

التقويم

الأفلاكي



صلاح فؤاد سليم



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿ وَقُلْ اَعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ وَرَسُولُهُ وَالْمُؤْمِنُونَ ﴾

صدق الله العظيم



التقويم النفسي



# التقويم النفسي

إعداد

صلاح فؤاد سليم

الطبعة الأولى

2006 م - 1425 هـ



مكتبة الخدمة العلم للنشر والتوزيع

سلام، فؤاد

العلوم النفسية / إعداد فؤاد صلاح سلام. - عمان: مكتبة  
المجتمع العربي، 2005.

( ) ص.

ر-: (2005/4/708).

الوصفات: المختبرات الشخصية // الإرشاد النفسي // علم النفس  
الفردية // الشخصية // علم النفس /

تم إعداد بيانات الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية

حقوق الطبع محفوظة للنشر

Copyright ©  
All Rights reserved

الطبعة الأولى

2006 م - 1426 هـ



مكتبة المجتمع العربي للنشر والتوزيع

عمان - شارع الملك حسين - مجمع الفعيفس التجاري

تلفص 4632739

ص.ب. 8244 عمان 11121 الأردن

# المحتويات

## الوحدة الأولى

- 9 ..... مبادئ القياس النفسي

## الوحدة الثانية

- 67 ..... اختبارات الذكاء العام

## الوحدة الثالثة

- 111 ..... اختبارات القدرة المتخصصة

## الوحدة الرابعة

- 141 ..... قياس الشخصية عن طريق الأحكام

## الوحدة الخامسة

- 157 ..... قياس الشخصية (وسائل التقرير الذاتي)

## الوحدة السادسة

- 183 ..... اختبارات التحصيل

- 215 ..... للمراجع





## مبادئ القياس النفسي





## مقدمة:

لواقع أن المعلومات التي تتوافر لنا عن المقاييس ذاتها: هل هو من نوع مقياس النسبة أو المسافة أو الرتبة، وعن موضوعيته ومدى تحرره من الأحكام الذاتية، وعن خصائصه: هل هو من النوع " الكلي " أو " التفاضلي "، وعن طبيعة البيانات التي يزودنا بها: هل هي تقديرات الآخرين أو استجابات مباشرة، وعن مدى تقبل شروط إصطلاحه، لهذه المعلومات جميعاً نقبلنا كثيراً في الحكم على أداة التقويم. إلا أن تقويمنا لأداة التقويم لا يكتمل إلا إذا توافرت لدينا مجموعة أخرى من المعلومات عن مدى دقة البيانات التي نحصل عليها بهذه الأداة، ومدى صلاحية الأداة ذاتها، وبعبارة أخرى ثابت الاختبار وصنقه.

## ثبات الاختبارات Test reliability:

يقصد بثبات الاختبارات، وأدوات التقويم علاقه أن تكون على درجة عالية من الدقة والإتقان والاتساق والاطراد فيما تزودنا به من بيانات عن سلوك المفحوص.

ويستخدم مفهوم ثبات بمعنىا الواسع ليدل على مدى اعتماد الفروق الفردية في درجات الاختبار على أعطاء الصدفة المتضمنة في القياس من ناحية وعلى الاختلافات الحقيقية الأممية في الصلة أو الخاصية الميكولوجية موضع الاهتمام من ناحية أخرى. ويمكن صياغة المشكلة صياغة فنية على النحو الآتي: ما هي النسبة في الثبات الكلي لدرجات الاختبار التي يمكن اعتبارها من

نوع تبين الخطأ error variance وبالطبع فإن ما نعتبره من نوع تبين الخطأ في موقف معين يمكن اعتباره تبيناً حقيقياً في موقف آخر. فضلاً إذا كان اهتمامنا متجهاً نحو دراسة تجنب المزاج فإن التغيرات اليومية في درجات اختبار بقيس الاثراح والاتقاض تصبح مرتبطة بفرض الاختبار نفسه وتكون بالتالي جزءاً من التباين الحقيقي في الدرجات. أما إذا كان الاختبار يهدف إلى قياس الخصائص الأكثر استقراراً في الشخصية تصبح هذه التغيرات اليومية بذاتها من نوع تبين الخطأ.

وعلى ذلك يمكننا أن نعرف تبين الخطأ بأنه يتعلق بالظروف والظروف التي لا تتصل من قريب أو بعيد بفرض الاختبار. ولذلك فعندما يحاول المدرس أو المرشد النفسي أو الأخصائي النفسي أن يحافظ على توحيد الظروف الاختبارية وتقديرها وذلك ببطء مكان الاختبار، وتعليماته وزمنه وعلاقته بالمختبرين وغيرها من العوامل فإنه بذلك يقلل من تبين الخطأ ويجعل درجة الاختبار أكثر ثباتاً واستقراراً. ومع ذلك فلا يمكن اعتبار الاختبار أداة ثباتاً كاملاً يصل إلى درجة ثبات مقاييس الظواهر الطبيعية (كالمتري للأطوال - الكيلو جرام للوزن). ولذلك لا بد أن يحدد صانعو الاختبارات مقدار ثباتها في دراسة التحليلات بحيث يكون دليلاً للأخصائيين عندما يطبقون هذه الاختبارات في ظروف معيارية مقننة.

وبالطبع يمكن أن توجد أنواع من ثبات الاختبارات بعدد الظروف التي يمكن أن تؤثر في درجات هذه الاختبارات ولا تكون كلها أو بعضها غير متعلقة بفرض الاختبار. إلا أن المصادر الرئيسية التي يمكن أن نرجع إليها تبين

الخطأ والتي تحدد للمقياس العامة لثبات الاختبارات يمكن تصديقها في ضوء الطرق المختلفة لصلب الثبات.

### (أ) إعادة الاختبار

من المصادر الأساسية لتباين الخطأ وما يتمثل في للتغيرات العشوائية التي تتنبأ الأفراد من جلسة اختبار إلى أخرى. وقد تنتج هذه التقلبات عن ظروف تطرأ على الموقف الاختباري ولا يمكن التحكم فيها مثل التغيرات الفجائية في الجو أو حدوث ضوضاء مفاجئة (وغيرها من مشتتات الانتباه). وقد تنتج عن تغيرات في ظروف المفحوص نفسه كما تتمثل في حالات المرض والتعب والاضطراب الانفعالي والقلق وغيرها. وهذه العوامل تدل على المدى الذي تتأثر به درجات الاختبار بالتقلبات اليومية العشوائية في المفحوص أو في ظروف للمواقف الاختباري.

ومن الواضح أن درجة الاستقرار الوتقي Temporal Stability تعتمد إلى حد ما على طول الفترة الزمنية التي تتضمنها والتي نقرر بعدها ثبات الاختبار. وتوجد أسئلة عديدة لاختبارات لها ثبات مرتفع إذا قدر هذا الثبات بعد فترات زمنية قصيرة (أيام فلال أو أسابيع فلال) ولكن ثباتها ينخفض لتخلفاً شديداً عندما تمت هذه الفترات إلى عشر سنوات أو خمسة عشر سنة. وكثير من اختبارات تكاء أطفال ما قبل المدرسة لها ثبات مرتفع في فترة ما قبل المدرسة ولكنها تفشل في التنبؤ بكاء الأطفال في مرحلة الطفولة المتأخرة أو للمراقبة أو للضح بسبب ما يطرأ على الطفل من تغيرات ارتقائية في النواحي العقلية.

ولا شك في أن التنبؤات ذات المدى القصير (والتي تحدث خلال فترات تتراوح بين ساعات قليلة وأشهر قليلة) يمكن اعتبارها من نوع تبين الفضا في درجة الاختيار. ولذلك فهي ممارسة العملية عندما تقوم بتخمين مثل هذا النوع من التنبؤات لا بد من أن تكون الفترة الزمنية قصيرة، وتكون أقصر عند اختبارنا للأطفال الصغار منها عند اختبارنا للكبار بسبب سرعة معدل النمو في الطفولة عنه في مراحل النمو الأخرى. وبالنسبة لأي نوع من المفحوصين يجب ألا تزيد فترة ما بين الاختبارين عن ستة أشهر بحال من الأحوال.

ومعنى ذلك أن التغيرات التي تحدث في الأداء النمائي للمفحوصين بعد فترات زمنية طويلة يمكن اعتبارها ضرباً من النمو والتقدم في الملوكة وليس مجرد تغير عشوائي، بل قد تنمي هذه التغيرات إلى ميدان من السلوك أوسع نطاقاً مما يتناوله الأداء الاختباري ذاته. ومن ذلك أن المفحوص قد يتغير أدائه في اختبارات الاستعداد المدرسي أو لفهم الميكانيكي أو الحكم اللغوي بعد فترة زمنية طويلة حتى سنوات بسبب تعرض المفحوص لخبرات غير عادية في هذه الفترة، بعارة أخرى يمكن للفرد أن يتغير مركزه النمائي في الاختبار بعد هذه الفترة صعوداً أو هبوطاً بالنسبة لأقرانه من نفس العمر بسبب ظروف خاصة تعود إلى المنزل أو المدرسة أو البيئة الاجتماعية أو بسبب المرض أو الاضطراب الانفعالي.. إلخ.

ويجب أن نشير إلى أن الوظائف النفسية ذاتها قد تختلف فيما بينها في مدى ما تظهره من نبضات وتقلبات وتقنية حسناً قد تتأثر وظيفة سلوكية مثل تزلزل الحركات الدقيقة للأصابع بأي تغير ولو كان طفيفاً في ظروف المفحوص بينما نجد أن سلوكاً آخر مثل الاستدلال لا يسهل التأثير فيه بالمؤثرات

الممارسة. ولذلك إذا أردنا أن نحصل على تقدير دقيق لاتزان الأصابع فإننا نحتاج إلى إعادة الاختبارات مرات عدة في أيام متوالية، بينما تكفي جلسة اختبار واحدة لقياس الاستدلال. ومعنى ذلك أنه من الضروري تحليل أهراس الاختبار وطبيعة السلوك الذي يدعى الاختبار أنه قادر على التنبؤ به تحليلاً دقيقاً وصيفاً قبل اختبار أنسب للطرق لحساب ثباته.

ويقدر معامل الثبات الذي يتأثر ببيان الخطأ الناتج عن التقلبات الزمنية بطريقة إعادة الاختبار بصورته ذاتها على نفس الأفراد في مناسبات مختلفة، ويسمى المعامل للناتج (وهو في حقيقته معامل ارتباط بين مجموعتي نتائج الاختبار في كل مناسبتين طبق فيهما) معامل الاستقرار Coefficient of stability. ورغم بساطة هذه الطريقة فإنها تكتسب بعض المشكلات، منها مثلاً أن الممارسة قد تنتج تحقراً من التحسين في درجات إعادة الاختبار. ومنها أيضاً أنه لو كانت الفترة بين مواقف الاختبار وإعلانه قصيرة فإن المفحوص قد يستدعي بعض استجاباته السابقة، أو بعارة أخرى قد يحصل على نفس النمط من الإجابة من حيث الوضع النسبي للصواب والخطأ ولكن بسبب الوظيفة السلوكية موضع القياس، وبالتالي لا تكون الدرجات التي حصلنا عليها من موقعي الاختبار ممثلة بعضها عن بعض كما أن معامل الارتباط بين مجموعتي النتائج قد يكون مرتفعاً ارتفاعاً زائفاً. وكذلك من مشكلات طريقة إعادة الاختبار أن طبيعة الاختبار ذاتها تتغير بسبب الإعادة وخاصة إذا كانت أسئلة الاختبار تتطلب استخدام الاستدلال أو للحق. ففي مثل هذه الأسئلة إذا قام المفحوص باكتساب العلاقة أو معرفة للمبدأ المتضمن في المسألة أو توصل إلى حل المشكلة، فإنه يمكن أن يستعد الإجابة الصحيحة في المستقبل دون المرور بخطوات التفكير للوسيلة وعلى ذلك فإن هذه الطريقة لا تصلح إلا مع

الاختبارات التي تتأثر كثيراً كبراً بالإعادة أو للتكرار مثل اختبارات التمييز الحسي والاختبارات الحركية. أما بالنسبة لمعظم الاختبارات النفسية فإن طريقة إعادة الاختبار ليست ملائمة لها.

#### (ب) الصور المتكاثفة:

من المألوف لدى الكثير من الطلاب أن يدخلوا الامتحان ' فيسعدهم الحظ ' بعدد كبير من الأسئلة يتناول الموضوعات التي اهتموا بدراستها ومدكرتها أكثر من غيرها، وقد يحدث العكس في كثير من الأحيان. هذا الموقف يدل على مصدر آخر من مصادر تباين الخطأ في درجات الاختبار وهو مدى اعتماد درجة المفحوص في الاختبار على عوامل تتعلق بطريقة اختبار الأسئلة. فإذا افترضنا أن قام باحث آخر بإعداد اختبار تتوافر فيه نفس مواصفات الاختبار الأول فإلى أي حد تختلف درجة المفحوص في الاختبارين ؟ الواقع أن الفرق في الدرجة التي يحصل عليها نفس الأفراد في هذين الاختبارين هو نوع الثبات التي يتأثر بعونة الأسئلة التي يتكون منها كل من الاختبارين. فمثلاً إذا افترضنا أن باحثاً قام بإعداد اختبار القياس للمعلومات الميكانيكية يتكون من 50 سؤالاً، ثم قام باحث آخر بإعداد اختبار مكافئ " أي يتكون من 50 سؤالاً كذلك نفس العرض "، ولنفرض أن أسئلة كل من الاختبارين قد صيغت بنفس الدرجة من الماهرة بحيث تتناول نفس المستوى من الصعوبة، فإنه تبعاً لمبادئ المصدقة للبحث في الخبرات الماضية للمفحوصين نجد أن للصعوبة النفسية لمجوعتي الأسئلة سوف تختلف من شخص لآخر، وعلى ذلك فإن الاختبار الأول قد يحتوي على عدد كبير من الأسئلة لم يلفه المفحوص (أ) في

حين تكون لغته بأسئلة الاختبار الظلي أكبر. وقد يحدث العكس بالنسبة للمفحوص (ب).

ويسمى معامل الثبات في هذه الحالة بمعامل التكافؤ أو التساوي Coefficient of equivalence، يدل على مدى الارتباط بين درجات اختبار، ومقياس آخر متكافئ معه يعطى في نفس الوقت. ويمكن أن نحصل عليه بإعطاء صورتين من الاختبار في وقت مقارب جداً، بشرط أن تكون هاتان الصورتان متماثلتين متكافئتين، أي تقيسان نفس الصفات للعلماء الأذنين بنفس الطريقة. ولكن، كما سبق أن أوضحنا، قد تؤثر الصفات الخاصة كثيراً مختلفاً على الاختبارين مثل المعلومات الخاصة التي قد تساعد في الإجابة على سؤال معين من أسئلة أحد الاختبارين ولكنها لا تساعد في الاختبار الآخر.

وتباين الخطأ المتضمن هنا يدل على التغيرات في الأداء من مجموعة معينة من الأسئلة إلى مجموعة أخرى، أي الخطأ الذي يرجع إلى عينة الأسئلة. ولكنه لا يتضمن تباين الخطأ الذي يعود إلى العوامل الوقتية، والذي سبق أن تحدثنا عنه، والذي يتمثل في التغيرات في الأداء من موقف اختبائي إلى آخر. ويمكن الجمع بين هذين النوعين من تباين الخطأ إذا قمنا بتطبيق الصورتين المتكافئتين من الاختبار في مناسبتين مختلفتين \* أي لا نطبقهما في نفس الوقت \*. ويمكن بهذه الطريقة تقادي بعض مشكلات طريقة إعداد الاختبار نفسه. ويكون معامل الثبات هو معامل الارتباط بين درجات صورتين الاختبار، وفي هذه الحالة يدل على كل من الاستقرار الوقتي واتساق أو إطراد الاستجابة لعينات مختلفة من الأسئلة. ويمكن أن نسمي هذا المعامل معامل الاستقرار والتكافؤ معاً.



وعند إعداد الصور المكافئة من الاختبارات يجب أن نتأكد من أن هذه الصور فعلاً متكافئة. والصور المكافئة هي في جوهرها اختبارات تم إعدادها كل على حدة بطريقة مستقلة بحيث تتوفر فيها نفس المواصفات، أي أن تحتوي على نفس العدد من الأسئلة، وأن تكون صياغة هذه الأسئلة ممتثلة، وأن تتضمن الأسئلة محتوى واحد، وأن تتبادل الأسئلة في مستويات الصعوبة، وأن تتفق الاختبارات في جميع المظاهر الأخرى مثل التعليمات والزمّن والأسئلة التوضيحية والشكل العام. ولا يمكن أن نغزى الفروق في الدرجات من اختبار إلى آخر إلى تباين الخطأ الذي يرجع إلى عينة محتوى الاختبار إلا إذا كانت صورتان متكافئتين أو متعلاتين حقاً.

ومن المؤكد أن حاجتنا إلى الصور المتكافئة من الاختبارات تتعدى مجرد تحديد ثبات الاختبار. فالصور البديلة تزيد كثيراً في الدراسات التي نجريها حول آثار هامل تجريبي متوسط معين على الأداء الاختباري، كما يفيد استخدام الصور البديلة في إعطاء وسيلة لخفض احتمال للتدريب أو الفش.

إلا أن أهم صعوبات هذه الطريقة تتمثل في عدم إمكان إعداد صور متكافئة لكثير من الاختبارات لأن ذلك يصاعف الجهد والوقت والتكلفة المبذولة في بناء الاختبار. ومن المعروف أن إعداد اختبار نفسي واحد يتطلب جهداً عظيماً، فما بالك بإعداد أكثر من اختبار لقياس نفس السمة.

### (ج) التجزئة النصفية:

بدل تجانس الأسئلة عن مدى تسليق وإطراد أداء المنحصر على جميع الأسئلة التي يتكون منها الاختبار. فإذا كان لدينا اختباران تتكون أسئلة الأول

منهما من عمليات ضرب فقط بينما تتكون أسئلة الاختبار الثاني من العمليات الحسابية الأساسية الأربعة فإن الاختبار الأول سوف يكون أكثر إتساقاً وإطراداً بين أسئلته من الاختبار الثاني، لأنه في الاختبار الثاني (وهو اختبار غير متجانس الأسئلة) قد يحسن بعض المفحوصين الأداء في عمليات الطرح مثلاً بينما يسهئون أداء باقي العمليات.

فلذا زاد حجم تجانس الأسئلة، كل يتكون الاختبار مثلاً من 40 سؤالاً كل عشرة منها تتناول مبادئ المفردات والعلاقات المكانية والاستدلال الحسابي والسرعة الإدراكية على التوالي فقد لا تكون هناك علاقة أو قد تكون العلاقة ضئيلة للغاية بين أداء المفحوص على هذه الأمثلة المختلفة من الأسئلة.

ومن ذلك نتبين أن درجة التخص في اختبار ما تكون أكثر وضوحاً عندما نحصل عليها من اختبارات متجانسة التكوين نسبياً. وتبرز مشكلة وضوح معنى للدرجة الاختبارية إذا عدنا إلى المثال الأخير الذي ذكرناه على الاختبار غير المتجانس. فإذا حصل المفحوصان (أ، ب) على درجة مقدارها 20 في هذا الاختبار فهل يمكننا أن نستنتج من ذلك أن أداءهما متساو تماماً في هذا الاختبار ؟ الإجابة بالنفي دون شك. فقد ينجح المفحوص (أ) مثلاً في الإجابة على أسئلة المفردات والسرعة الإدراكية ويفشل تماماً في الإجابة على أسئلة الاستدلال الحسابي والعلاقات المكانية. أما المفحوص (ب) فقد يحصل على درجته من الإجابة على خمسة أسئلة السرعة الإدراكية و خمسة أسئلة للعلاقات المكانية و عشرة أسئلة للاستدلال الحسابي ويفشل تماماً في الإجابة على أسئلة المفردات. وهناك ما لا حصر له من الاحتمالات في مثل هذا النوع من الاختبارات غير متجانسة التكوين. ومعنى ذلك أن درجة المفحوص في مثل هذا الاختبار لها

معان مختلفة، بينما يكون معناها أكثر وضوحاً في الاختبارات المتجانسة (أو على حد تعبير ثرسون الاختبارات ذات التكوين للعملية البسيط).

ورغم أن الدرجة 20 في اختبار متجانس الأسئلة (معاني المفردات مثلاً) قد لا يكون لها معنى واحد تماماً في جميع الحالات، ولذا قد تنفي أن المفردات لأجل إجابة صحيحة على المفردات العشرين الأولى تقريباً إذا كانت هذه المفردات مرتبة درجة الصعوبة، وربما قد يترك كلمتين أو ثلاث من الكلمات السهلة أو قد يفضل في الإجابة عليها، وقد يجمع في الإجابة على كلمتين أو ثلاث أصعب أو أبعد من الكلمة العشرين في ترتيب الصعوبة. ومع ذلك فإن هذه الاختلافات القدرية قليلة وحشة إذا قررت بذلك فتي تلاحظها في الاختبارات الأقل تجانساً.

وهكذا بعد عدم تجانس أسئلة الاختبار مصدراً لتباين الخطأ، وبالتالي يؤثر في ثبات الاختبار. ولكن هل يختلف تجانس الأسئلة كمصدر لتباين الخطأ عن المصدر الذي سبق للحديث عنه وهو ملائمة عينة المحتوى عموماً ؟ الواقع أن الفرق بين هذين المصدرين يمكن أن يتضح إذا أعطينا مثلاً متطريلاً. لنفترض أن هناك اختباراً بقياس كل سؤال فيه عملية عقلية أو وظيفة نفسية مختلفة ولا يرتبط أحى ارتباطاً بالأسئلة الأخرى التي بقياس كل منها عمليات أو وظائف أخرى. ولا شك أنه في مثل هذه الحالة يمكن إعداد صورة متكافئة من هذا الاختبار يحتوي على نفس أنواع الأسئلة وتوزيعها على الصور الموجود في الصورة الأولى. وعندما نحطى الصورتين المتكافئتين قد تتفق درجاتهما اتفاقاً كاملاً مما يدل على الوجهة النظرية على معامل ثبات مرتفع لا يتأثر كثيراً بحجم الأسئلة كمصدر لتباين الخطأ. ولكن عندما نهتم بثبات في ضوء تجانس الأسئلة

نجد أن هذا المعامل يصل إلى الصفر أو قرب إليه، حيث أن إطراد الأداء من سؤال إلى آخر في أي من صورتَي الاختبار قد لا يكون له دلالة الفصل من مستوى الصنف.

ويسمى معامل الثبات الذي يتأثر بتجانس الأسئلة معامل الاتساق الدلخي، ويمكن الحصول عليه من تطبيق واحد لصورة واحدة من الاختبار، ثم اللجوء إلى ما يسمى طريقة التجزئة للنصفية. وبها يمكن الحصول على درجتين لكل فرد وذلك بتقسيم كل اختبار إلى نصفين يمكن المقارنة بينهما ثم حساب معامل الارتباط بينهما. ومن الواضح أن حساب الثبات بالتجزئة للنصفية أو بحارة أخرى معامل الاتساق الدلخي، لا يعطي مقياساً للتجانس الكلي للمقياس لأنه يقسم أسئلة الاختبار إلى نصفين ولا تتناول هذه الأسئلة سوياً سوياً.

وإختر المشكلات التي تراجها طريقة للتجزئة للنصفية هي كيف نجري الاختبار للحصول على أفضل قسمين للمقارنة. هناك عدة طرق لقسم الاختبار وخاصة إذا علمنا أن معظم الاختبارات لا يمكن تجزئتها إلى نصفين صالحين للمقارنة بينهما بمجرد تتبع التوصل الرسمي للأسئلة، وذلك بسبب الفروق في طبيعة الأسئلة ومستوى صعوبتها بالإضافة إلى الآثار التراكمية للتوصل في الأداء أو العمل والممارسة والتعب والمال وغيرها من العوامل التي تختلف في بداية الاختبار عنها في نهايته. ولذلك فإن إحدى الطرق الدقيقة والموضوعية للحصول على قسمين متماثلين في الاختبار هي تحديد مستوى صعوبة كل سؤال وذلك بليجاد النسب المئوية للأفراد الذين أجابوا عليه إجابة صحيحة، ثم توزع الأسئلة على قسمي الاختبار تبعاً لتكافؤ معنوي الصعوبة ولتشابه التضمين. ثم نطبق الاختبار على مجموعة أخرى من المفحوصين غير تلك

التي طبقنا عليها الاختبار لحساب مستويات الصعوبة، ونحصل على درجات الأفراد لتقدير الثبات أو معامل الاتساق للدخلي.

ومن الطرق الصالحة لمعظم الأغراض وفي نفس الوقت لا تمتزق وفقاً طويلاً ولا تتطلب جهداً زائداً طريقة تجزئة الاختبار إلى نصفين أحدهما يضم الأسئلة ذات الأرقام الزوجية والآخر يضم الأسئلة ذات الأرقام الفردية، وبخلفة إذا كانت الأسئلة مرتبة في الاختبار تبعاً لمستويات الصعوبة، إلا أننا يجب أن نحذر أنه في حالات مجموعات الأسئلة التي تتعلق بمشكلة واحدة مثل الأسئلة التي تتلو فترة محبة في اختبار للقراءة فإن هذه الأسئلة جميعها يجب أن نوضع في جزء من الاختبار لأننا لو قسمنا مجموعة الأسئلة من هذا النوع على جزئي الاختبار فإن التشبه بين الجزئين سوف يزداد زيادة مسطحة أو زائفة حيث أن أي خطأ في فهم المشكلة قد يؤثر في أسئلة نصفي الاختبار.

وقد أوضحنا أننا نحصل على معامل الاتساق للدخلي بحساب معامل الارتباط بين درجات نصفي الاختبار ولكننا يجب أن نشير إلى أن مثل هذا المعامل يعطينا ثبات نصف الاختبار فقط وليس الاختبار كله. وهذه الخاصية لا توجد بالطبع في الطرق المباشرة كإعادة الاختبار أو الصور المتكافئة. ونحن نعلم أنه إذا تساوت الظروف الأخرى فإن الاختبار الأطول من حيث عدد أسئلته يكون أعلى ثباتاً، لأننا لو حصلنا على عينة أكبر من الملوكة نصل إلى مقاييس أكثر دقة وأكثر استقراراً ويمكن تقدير الأثر الذي تحدثه إطالة الاختبار أو اختصاره على معامل الثبات باستخدام معادلة سبيرمان برلون المشهورة وهي:

$$r = \frac{r_{12}}{1 + (n-1)r_{11}}$$

حيث أن  $r$  = معامل الثبات الذي نريد الحصول عليه.

رأ = معامل الثبات الذي حصلنا عليه.

ن = عدد مرات إطلاق الاختبار أو اختصاره أي أنه إذا زاد عدد الأسئلة من 25 - 100 سؤال فإن (ن) تكون 4 أما إذا اختصر من 60-30 سؤال فإن (ن) تصبح  $\frac{1}{2}$  . ومعنى ذلك أن هذه المعادلة حينما تطبق على معامل الثبات المقدر بطريقة التجزئة النسبية فإنها تتضمن إطلاق الاختبار إلى النصف أي تصبح المعادلة على النحو الآتي:

$$r = \frac{r_a 2}{r_a + 1}$$

إلا أن معادلة سبيرمان براون تفترض أن للفحريات في درجات نصفي الاختبار متساوية تماماً وهو افتراض يصعب في كثير من الأحوال تحقيقه حتى ولو بدأ نصف الاختبار على درجة كبيرة من التكافؤ. ولذلك فقد يستعاض عن ذلك باستخدام معادلة جتمان (12) التي تستبدل هذا الافتراض ويمكن صياغتها على النحو الآتي:

$$r_a = 1 - \left[ \frac{1 + \frac{\sigma^2_b + \sigma^2_e}{\sigma^2_e}}{2} \right]$$

وفي هذه المعادلة نجد أن ع أو ع ب = الانحرافات المعيارية لتصني الاختبار.

ع ك = الانحراف المعياري للدرجات الكلية  
في الاختبار.

ومن الواضح أن هذا الإجراء لا يتطلب حساب معامل الارتباط

(د) تطويل التقييم

تصلح طريقة التجزئة النصفية لحساب معامل الاتساق الداخلي (تجانس نصفي الاختبار)، أما بالنسبة لموضوع تجانس الأسئلة ذاتها فإن بعض العلماء يرون أنه لا يدخل في نطاق مناقشة موضوع ثبات الاختبارات وإنما هو من خصائص الاختبارات الجيدة ولا يمكن إدراجه تحت أي من المفاهيم التقليدية للثبات. ومع ذلك فلا يمكن استبعاده تماماً من مناقشة الثبات حيث أنه يتعلق بمدى استقرار استجابات المفحوص. على أسئلة الاختبار واحداً بعد الآخر. وعلى كل فإن مفهوم لتجانس الكلي يجب أن يتميز تمييزاً واضحاً عن غيره من صور الثبات التي سبق لنا مناقشتها. وعادة ما يسمى "معامل الاتساق داخل أسئلة الاختبار" *Interitem Consistency* وهو مصطلح آخر قد لا يعطي نفس ما يعنيه مصطلح معامل الاتساق الداخلي. فمعامل الاتساق داخل الأسئلة يشير إلى التجانس الكلي للاختبار بالمعنى المباشر، أما معامل الاتساق الداخلي فهو أقرب إلى معنى التكاثر كما يتجلى في الصور المتكافئة للاختبار. ومعنى ذلك أنه قد يوجد لدينا اختباران لهما معاملات ثبات مرتفعة حسب إما بالصور المتكافئة أو بالتجزئة النصفية ولكنهما قد يختلفان في المعاملات الخاصة بالاتساق داخلي الأسئلة إذا اختلفا في درجة تجانس الأسئلة التي يتكون منها صورنا

الاختبار أو نصفاء. والواقع أن الفرق بين معامل الاتساق الداخلي (الذي يحسب بالتجزئة النصفية كما سبق أن أوضحنا) وبين معامل الاتساق داخل الأسئلة يمكن اختباره دليلاً على مدى التجانس أو التباين في أسئلة الاختبار.

وأشهر طرق لحساب معامل اتساق الأسئلة طريقة كيردر ريتشاردسون وهذه الطريقة تشابه مع طريقة حساب معامل الاتساق الداخلي في أنها تستخدم إجراء اختبارياً واحداً للاختبار الواحد، ولكنها بدلاً من أن تتطلب تجزئة الاختبار إلى نصفين متكافئين فإنها تعتمد على فحص أداء الأفراد على كل سؤال على حدة ولذلك فهي تنتمي إلى طرق تحليل التباين بالمعنى العام، وأشهر المعادلات الإحصائية التي تستخدم في هذا الصدد هي:

$$r = \frac{c^2 - \text{مجموع } x}{c^2} \left( \frac{n}{1 - n} \right)$$

حيث أن  $r$  = معامل ثبات الاختبار.

$n$  = عدد الأسئلة التي يتكون منها الاختبار.

$c$  = الانحراف المعياري لدرجات الاختبار.

مجموع  $x$  = يحسب بإعداد نسبة الأفراد الذين أصابوا في إجاباتهم على كل سؤال ونسبة الذين أخطأوا ثم لضرب النسبتين في بعضهما ثم نجمع حواصل الضرب بالنسبة لجميع الأسئلة.



ومن الواضح أن معامل الاتساق داخل الأسئلة كما يحسب بطريقة كيوودر ريتشاردسون هو في الواقع متوسط معاملات الاتساق الداخلي الناتجة من عدد من التجزئات للاختبار. أما معامل الاتساق الداخلي فيعتمد على نجدة تهدف إلى الحصول على مجموعتين متساويتين أو متكافئتين من الأسئلة. وعلى ذلك إذا لم تكن أسئلة الاختبار على درجة كبيرة من التجانس فإن معامل الاتساق داخل الأسئلة يصبح أقل من معامل الاتساق الداخلي.

#### (هـ) الممتحنون والمصححون:

بعد أن نستبعد الأنواع الثلاثة السابقة من ثباين الخطأ من الثباين الكلي (الدرجة التي يحصل عليها المصححون في الاختبار) فإن الثباين المتبقى يرجع إلى الفروق الفردية للخصائص موضوع الاهتمام، وبعضه الآخر إلى عوامل أخرى غير متعلقة ولكن يمكن ضبطها تجريبياً. فمثلاً لا يسجل الباحثون أخطاء القياس الناتجة عن تطبيق الاختبار في ظروف مشتتة للانتباه، أو في زمن أطول أو أقصر مما يحدده كتيب الاختبار، أو بإعطاء تعليمات مختلفة عن التعليمات المحددة لأن هذه الأخطاء جميعاً يمكن استبعادها والتغلب عليها في الموقف الاختباري نفسه. ولذلك فالحسن لا نجد في العادة معاملات ثبات خاصة تشير إلى " ثباين الخاص بثقت الانتباه " أو " ثباين اختلاف الزمن " أو " ثباين اختلاف التعليمات " ... إلخ. ثم إننا نجد أن معظم الاختبارات تحدد إجراءات مقبنة للتطبيق والتسميحي بحيث نفترض مع وجود هذه الإجراءات والالتزم بها أن " ثبات الممتحن " و " ثبات المصحح " يصل إلى درجة كبيرة من الكفاءة تجعل الاختبار مفيداً في الأغراض التطبيقية والعملية وبالتالي لا يكون في حاجة إلى تعديل مثل هذا النوع من الثبات. ويصدق هذا القول على وجه الخصوص

بالنسبة للاختبارات الجماعية التي تطبقها على جماعات كبيرة وتصحبها الآلات الحاسبة الإلكترونية، أي أنه في هذه الحالات جميعاً لا يتطلب الأمر إلا التأكد من التتابع المنطقي لإجراءات التطبيق والتصحيح وتكون المشكلة في الضبط التجريبي لظروف إجراء الاختبار.

إلا أننا نلاحظ أنه في الاختبارات الفردية مثل متلفرد يبنيه يلعب الممتحن دوراً أكثر تعقيداً، لأن إجراءات الاختبارات في هذه الأحوال ليست مقلنة تقنياً جامداً بل يعتمد معظمها على نجاح الممتحن في تكوين العلاقة بينه وبين المفحوص، وعلى دوافع المفحوص. وقد يتطلب الأمر أن يقوم الممتحن بتقويم أداء المفحوص خلال عملية تطبيق الاختبار لأن مثل هذا الأداء يحدد كيف يمتحن الممتحن في موقف الاختبار. ولذلك فقد يختلف الممتحنون في النتائج التي يحصلون عليها من نفس المفحوصين حتى ولو توفرت في هؤلاء الممتحنين أفضل الشروط وهذا الاختلاف قد يكون من نوع تبليغ الخطأ الذي يمكن أن ترجعه إلى الفروق الفردية بين الممتحنين ويؤثر بالطبع في النتائج الكلية لدرجات المفحوصين. ولذلك فإذا كانت الاختبارات التي تستخدمها من النوع الذي تلعب فيه فردية الممتحن دوراً هاماً، أو إذا كانت أدوات التقويم من الذي يتأثر بالعوامل الذاتية مثل المقابلة الشخصية، تكون في حاجة ماسة إلى تقدير " ثبات الممتحن " وخاصة إذا كان من الضروري معرفة العلاقة بين تقديرات أكثر من ممتحن لنفس الفرد أو الأفراد.

وبالمثل نجد أن بعض الاختبارات تثير مشكلة " ثبات المصحح Reliability of Scorer وخاصة لاختبارات المقال والاختبارات الإسقاطية

بروجه عام، وهي الأدوات التي تسمح بمقدار كبير من التفسير الذاتي من جانب المصحح وهو الممتحن أيضاً في كثير من الأحوال.

ومن الواضح أن معامل ثبات المصحح يمكن الحصول عليه بحساب معامل الارتباط بين الدرجات التي يعطيها مصممان أو أكثر لنفس الاختبارات أو نفس الأفراد. بعبارة أخرى فإن كل مقياس يحمل على درجتين أو أكثر من تصحيح اختبار واحد. وهذه الطريقة لا تتضمن بالطبع توافقي القصور الموجودة في طريقة إعادة الاختبار. أما معامل ثبات الممتحن فهو أكثر صعوبة لأنه بطبيعته يتضمن يتضمن إعادة اختبار، وبالتالي فإن قوام ممتحن بإعادة الاختبار يعني الوقوع في جوانب القصور التي سبق الحديث عنها في طريقة إعادة الاختبار التي يقوم بها نفس الممتحنين. بالإضافة إلى أن معامل الارتباط الناتج يعكس عوامل الاستقرار لوقت وثبات الممتحن معاً. ويمكن اللجوء إلى نمط مختلف من التخطيط التجريبي يعتمد على تحليل التباين لمزل تباين الخطأ الذي يمكن إرجاعه إلى الممتحنين. وعلى كل فيمكن أن نعتبر كلاً من معامل ثبات المصحح ومعامل ثبات الممتحن من نوع واحد ونطلق عليه اسم معامل الموضوعية Coefficient of objectivity.

ويمكن أن نلخص أنواع معاملات الثبات وطرق تقديرها في الجدول التالي.

جدول رقم (1)

المعامل	طريقة تقديره	تباين الخطأ
معامل الاستقرار	إعادة الاختبار بنفس صورته ونفس الممتحنين في مناسبة مختلفة	التغيرات المؤقتة
معامل التكاليف	تطبيق الاختبار مع صورته متكافئة له في نفس الوقت	عينة محتوى الاختبار
معامل الاستقرار والتكاليف معاً	تطبيق الاختبار ثم تطبيق صورة متكافئة له في مناسبة مختلفة	التغيرات المؤقتة عينة المحتوى
معامل الاتساق الداخلي	للجزئة النصفية	عدم نجاة نصف الاختبار
معامل الاتساق الدخلى	تحليل المفردات باستخدام طريقة كيوود ريتشاردسون	للتجارب الكلية لأسئلة الاختبار
معامل الموضوعية	تكرار المصحح وتكرار الممتحن	ذاتية الممتحن، التغيرات المؤقتة

## صدق الاختبارات

يتعلق موضوع صدق الاختبارات بما يقوسه الاختبار وإلى أي حد ينجح في قياسه، وهو موضوع لا يقتصر على عملية القياس وإنما قد يمتد إلى المنهج التجريبي بصفة عامة والذي يتطلب اختبار الفروض العلمية للتحقق من صحتها.

وإذا كان الصدق مسألة على درجة كبيرة من الأهمية في تقرير صلاحية أو فائدة أي فرض علمي أو في تحديد معنى مفهوم معين فليج في الواقع هو الذي يحدد قيمة الاختبار ومغزاه، ويجب أن نقرر منذ البداية أن عنوان الاختبار لا يدل بالضرورة على ما يقوسه الاختبار لأنه عادة ما يكون مجرد كلمة أو عبارة قصيرة تهدف إلى التعرف على الاختبار. ولا يمكن أن نستدل على الوظيفة السلوكية التي يتبناها الاختبار إلا بفحص مجموعة من السمكات (أو غيرها من المصادر الموضوعية للمعلومات) المتعلقة نسبياً عن الاختبار.

ويمكن القول بوجه عام أن جميع طرق تحديد صدق الاختبارات تبحث عن العلاقة التي تربط أداء الأفراد في الاختبار بحقائق ومعلومات خاصة بالوظيفة السلوكية موضوع الاختبار يمكن ملاحظتها ملاحظة مستقلة. وهناك كثير من الطرق التي تستخدم في هذا الصدد ولها مسميات كثيرة، إلا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التي وصفها كتيب التوصيات لفدية للجمعية الأمريكية لعلم النفس، والتي عالجتها أنستازي وكرونباك بالتفصيل.

### أ- صدق المحتوى أو المضمون:

يقصد بصدق محتوى الاختبار أن نقوم بفحص مضمون الاختبار فحصاً دقيقاً منتظماً لتحديد ما إذا كان يشمل على عينة ممثلة لميدان السلوك الذي يقوسه.

ومثل هذا النوع من الصدق نستخدمه في العادة في تقويم الاختبارات التحصيلية التي نعدّها بقياس مقدار ما تمكن التلاميذ من إتقانه من جوانب مهارة معينة أو ما حصلوه من مقرر دراسي أو برنامج مهني. وقد يبدو أن مجرد فحص مضمون الاختبار يكفي للتأليل على صدقه وصلاحيته للغرض الذي وضع من أجله، فاختبار معين في الضرب أو التاريخ قد يبدو صادقاً بالتحريف إذا اشتمل على أسئلة تدور حول عملية الضرب أو التاريخ.

ولكن الأمر ليس على هذا القدر من السهولة لأن من أهم المشكلات التي درلهمها هنا هي مشكلة عبء المحتوى. فميدان المحتوى الذي نخبره يجب أن يتم تحليله تحليلاً دقيقاً منتظماً حتى نتأكد من أن جميع الجوانب الأساسية يمكن أن تغطيها أسئلة الاختبار تغطية ملائمة وبالنسب المناسبة حتى لا يكون لدينا عدد كبير من الأسئلة في جانب معين على حساب الجوانب الأخرى. وهذا يتطلب بداهة ذي بده وصفاً تفصيلياً دقيقاً للمحتوى أو المضمون الذي نعد له الاختبار التحصيلي، كما يتطلب شعول الاختبار لأهداف التدريس أو للتدريب وليس مجرد المادة الدراسية أو المقررات. وبذلك يصبح صدق المحتوى أكثر اعتماداً على مدى الاتصال بين استجابات الفرد على أسئلة الاختبار وبين ميدان السلوك الذي نخبره.

وتوجد عدة طرق تجريبية للتحقق من صدق المحتوى. فمثلاً إذا أمكن إصدار صورتين متكافئتين من الاختبار فيمكن أن نطبق إحداهما قبل التدريب والأخرى بعده للتحقق من حدوث أي تحسن في الدرجات. وهناك طرق أخرى مثل دراسة أنماط الأخطاء الشائعة في الاختبار، وتحليل طرق العمل التي يستخدمها المفحوصون وذلك بإعطاء الاختبار فردياً مع توجيههم إلى " التفكير

بموت مرتفع " لثناء حل مشكلات الاختبار. كما أن أثر السرعة يمكن تحديده بمعرفة عدد الذين فشلوا في الانتهاء من الاختبار أو بأي طريقة أخرى من الطرق الملائمة.

وعندما يكون الهدف أن نكتشف أثر القدرة على قراءة التعليمات في الأداء الاختباري (وخاصة إذا كانت تلك القدرة غير مرتبطة بهذا الأداء) فيمكن أن نحسب معامل الارتباط بين درجات الاختبار التجريبي ودرجات اختبار مقس في فهم القراءة. أما إذا كان الاختبار للتجربي يهدف إلى قياس القدرة على فهم القراءة فيمكننا أن نعطي للكرد الأسئلة التي يحتويها الاختبار دون أن نعطيهم للنس القرائي الذي تعتمد عليه لمعرفة عدد الأسئلة التي يمكنهم أن يجيبوا عليه معتمدين على معلوماتهم السابقة، إلى غير ذلك من العوامل التي قد تؤثر في مثل هذه الاختبارات.

وكما سبق أن قلنا أن صدق المحتوى يفيد في مراجعة اختبارات التحصيل ولكنه لا يكون كافياً في اختبار القدرات والاستعدادات ومقاييس الشخصية، بل قد يكون مضللاً في كثير من الأحيان. فطبي القوم من أن هذه الاختبارات لا بد أن يتولر فيها محكا الملائمة والفاضية بالنسبة لمضامينها وخاصة في المراحل الأولى لإعدادها، فإلغا فيما بعد للحا إلى وسائل التحقق التجريبي المضبوط التي سوف نتحدث عنها فيما بعد لتقدير صدقها. ويتضح ذلك إذا علمنا أن لاختبارات القدرات والاستعدادات والشخصية على عكس الاختبارات التحصيلية لا تتشابه تشابهاً داخلياً كثيراً مع ميدان السلوك الذي نعد هذه الاختبارات عينة منه، وبذلك لا يكون لمضامينها أي دور في الكشف عن القروض التي أنت بمؤلف الاختبار

إلى انتقاء مضمون معين يستخدمه في قياس صفة معينة، لأن هذه الغرض تتطلب تحقيقاً تجريبياً حتى يمكن تحديد صدق الاختبار.

كما أن اختبارات القدرات والشخصية على عكس الاختبارات التحصيلية لا تعتمد على مقرر دراسي أو على برنامج تدريبي أو على مجموعة موحدة من الخبرات السابقة يمكن منها صياغة مضمون الاختبار، وإنما يتفاوت الأفراد كثيراً في طرق عملهم وفي العمليات النفسية التي يستخدمونها عند الإجابة على نفس الأسئلة، وعلى ذلك فإن الاختبار الواحد قد يقيس وظائف سلوكية مختلفة في الأفراد المختلفين. وفي مثل هذه الأحوال يستحيل علينا تحديد لوظائف السلوكية التي يقيسها الاختبار من مجرد فحص معتاد قد يحل طالب للامتحان مشكلة معينة مستخدماً للتعبير اللغوي (أو لرمزي عموماً ومنه الرياضي)، بينما يحل الميكانيكي نفس المشكلة ويصل إليه، نفس الحل مستخدماً للتصور البصري المكاني. كما أن اختباراً معيناً للاستدلال الحسابي لطلاب المرحلة الإعدادية قد يقيس مجرد الفروق الفردية في سرعة العمليات العددية عند طلاب الجامع.

ويجب ألا نخلط بين صدق المحتوى والصدق الظاهري. والواقع أن الصدق الظاهري ليس صدقاً بالمعنى الطبي للكلمة لأنه يدل على ما يبدو أن الاختبار يقيس من الظاهر لا على ما يقيسه الاختبار بالفعل. كما أنه يعني ما إذا كان الاختبار يبدو صادقاً في نظر المفحوصين من ناحية، وفي نظر الفاحصين، سواء أعدوا إعداداً قديماً لم يمدوا، من ناحية أخرى. والواقع أن موضوع الصدق الظاهري يدخل في باب العلاقة الطبية التي يكونها الفاحص مع المفحوص، وهو من الأمور المرغوب فيها دون شك رغم ما يسببه من خلط



وتدخل بينه وبين غيره من مفاهيم الصدق الحقيقية. فوما لا شك فيه أنه إذا بدأ الاختبار للفاحصين أو المفحوصين لا معنى له، أو لا علاقة له بالظاهرة موضوع القياس، أو غير مناسب أو سانجاً أو مخيفاً، فمن المضمّن أن تعاون هؤلاء سيكون ضعيفاً حتى ولو كان الصدق الحقيقي للاختبار عالياً. عبارة أخرى لا بد أن تتوفر في الاختبار ما يسمى الصدق الظاهري حتى يكون أكثر فاعلية في موقف القياس العملية.

#### ب- الصدق التقني:

يصد بالصدق التقني قدرة الاختبار وفاعليته في التنبؤ بنتيجة معينة في المستقبل، ولهذا يفرم على المقارنة بين درجات المفحوصين في الاختبار ودرجاتهم في مقياس آخر مباشر للأداء اللاحق للفرد وهذا المقياس المحدد يسمى من الناحية الفنية " المحك أو الميزان. وهذا النوع من الصدق يتصل اتصالاً وثيقاً بالاختبارات التي تستخدمها في أغراض الانتقاء والتصنيف والتصنيف سواء في الأغراض التربوية أو المهنية أو العسكرية، وفي تحديد المرضي للمصابين الذين يحتمل أن يفيئوا من برنامج علاجي معين. وفي جميع هذه الأغراض تتشابه إجراءات الصدق، حيث تعطى الاختبار لمجموعة من الأفراد تنتمي للمجتمع الأصلي (موضوع الاهتمام)، ولكن الدرجات التي يحصل عليها أفراد هذه العينة لا تستخدم في تغذية أية قرارات خاصة بهم. كما يجب ألا تعطى هذه الدرجات إلى أي شخص أو سلطة يمكن له أو لها أن تتخذ قرارات تؤثر فيهم، حتى لا يطرأ على المحك أي تشويه أو خلط. وبعد ذلك تتم متابعة العينة الأصلية حتى تتوفر لدينا عليها المعلومات التي نستقيها من المحك والتي تؤدي بنا إلى تحديد

المدى الذي تتفق فيه درجات الاختبار مع النتائج التي نلاحظها كما تتمثل في درجات المحك، وهذا الاتفاق يحدد لنا مدى تكيفية الاختبار على أسس إحصائية.

وعادة ما يتضمن دليل الاختبار فهماً خاصة بالصدق للتنبؤ مشتقة من استخدام محكات مختلفة بهدف مساعدة من يستخدم الاختبار في فهم ما يقسمه وذلك على الرغم من أن مستخدمي الاختبارات قد لا يهتمون اهتماماً مباشراً بالتنبؤ بأي من المحكات المستخدمة. ومعظم المحكات التي نذكرها ككليات الاختبارات عادة ما تكون للحصول للمدرسي الأكاديمي العام، أو أداء الأفراد في برنامج تدريبي متخصص، أو أداء الأفراد في ميدان للعمل أو التطوير نفسه. ونفصل الحديث عن هذه المحكات كما يلي:

## 1- الحصول للمدرسي الأكاديمي العام:

هو أشهر هذه المحكات جميعاً وخلصه بالنسبة لاختبارات الذكاء. وإننا نرى هذه الاختبارات عادة بأنها مقاييس الاستعداد المدرسي. أما للمؤشرات التي تلجأ إليها كمقياس لمحك الحصول الأكاديمي فغالباً ما تكون للدرجات أو التقديرات المدرسية، ودرجات الاختبارات التحصيلية، ونتائج النقل والتخرج، ودرجات الشرف والتخصص، والجوائز والمنح، وتقديرات المعلمين لذكاء تلاميذهم، هذه المؤشرات المختلفة تعطينا معلومات خاصة بهذا المحك في جميع المستويات التعليمية من المرحلة الابتدائية حتى الكليات الجامعية والمعاهد العليا. وهذا المحك نستخدمه كما قلنا في تقدير صدق اختبارات الذكاء، وكذلك بطاريات الاستعدادات للفرقة، واختبارات القدرات العقلية والخاصة، وبعض اختبارات الشخصية. وعند حساب صدق الاختبار بفرض استخدامه في عمليات الانتقاء لأكاديمية معينة فمن الشائع أن نستخدم محكاً معيناً هو متوسط

درجات كل طالب مستجد. هذا المتوسط هو بدوره متوسط درجات الطالب في جميع المقررات التي يدرسها في السنة الأولى أو في السنة الإعدادية بالجامعة.

## 2- الأداء في برنامج تدريبي أو تعليمي معين:

وإذا كان ما نقوم به هو إعداد بطاريات الاستعدادات الفارقة والقدرات الفارقة فإن المحك الذي نستخدمه هو أداء الأفراد في برنامج معين من التدريب للمختصين. ففي حساب صدق اختبار معين للقدرة الميكانيكية يمكن الاعتماد على التحصيل النهائي في مقرر معين ينطق بإحدى المهن الصناعية. وكذلك يمكن أن نعتبر التحصيل في مقررات التطعيم التجاري مثل الاختزال والألة الكتيبة وسلك الفاتر، محكات لاختبارات الاستعدادات في هذه الميادين. وكذلك فإن أداء طلاب المعاهد الموسيقية والفنية قد يعتبر محكاً لحساب صدق اختبارات الاستعدادات للموسيقية والفنية. وقد حسب صدق معظم اختبارات الاستعدادات المهنية باستخدام محكات التحصيل في كليات الحقوق والطب والصيدلة والهندسة وغيرها. أما إذا كانت الاختبارات من النوع الذي نستخدمه في أغراض برنامج معين للانتقاء. فقد نستخدم سجلات تدريب كمصدر للمستويات التي نحتاج إليها. كما هو الحال في حساب صدق اختبارات انتقاء الطيارين حيث يكون المحك هو أداء الطيارين في موقف التدريب الفعلي.

ومن المؤشرات الخاصة التي يتضمنها محك الأداء في برامج التدريب يمكن أن نذكر الاختبارات التحصيلية التي نطبقها عقب الانتهاء من برنامج التدريب أو تدريبات المدربين أو النجاح والفشل في التدريب (الاستمرار والإكمال في مقابل الانقطاع والاستعدادات). وقد نستخدم درجات المدرسة الثانوية أو الكليات الجامعية والمعاهد العليا في مقررات معينة في تحديد صدق

بطاريات الاستعدادات، فمثلاً يمكن مقارنة درجات الطلاب في اختبار الفهم  
اللفظي، بدرجاتهم في مقررات اللغة العربية، ودرجات التصور البصري للمكاني  
بدرجات مقرر اللغة للعربية، ودرجات التصور البصري للمكاني بدرجات مقرر  
الهندسة... إلخ.

ولا بد أن يميز في هذا الصدد بين ما يسمى المحكات المتوسطة  
والمحكات القصوى أو النهائية بالنسبة لمؤشرات سجلات التكريب. فعدد إعداد  
اختبار لقياس الاستعدادات للطيران أو لقياس الاستعداد الطبي مثلاً نجد أن  
المحكات النهائية هي الأداء الفعلي في ميدان القتال، أو النجاح في ممارسة  
الطب. ويتطلب ذلك دون شك وقتاً طويلاً حتى يمكننا الحصول على معلومات  
من هذه المحكات، بل أنه من المشكوك فيه إمكان وصولنا إلى محك نهائي في  
الواقع بالإضافة إلى أنه، حتى إذا توفر لنا مثل هذا المحك النهائي، فإنه قد  
يخضع لعوامل كثيرة يصعب التحكم فيها مما يجعل المحك في النهاية لا فائدة  
منه. فمن الصعب مثلاً أن تقوم درجة نجاح التلميذ للأطباء الذين يمارسون  
مختلف التخصصات في أجزاء مختلفة من الوطن ولهذا السبب تلجأ في العادة  
إلى المحكات المتوسطة التي ذكرنا أمثلة لها فيما سبق مثل سجلات الأداء في  
مرحلة معينة من مراحل التكريب.

### 3- الأداء على العمل نفسه:

هو أكثر المحكات صلاحية لأنه يستند على سجلات تتبعية لأداء الواقعي  
والفعلي للأفراد في العمل. وقد استخدم العلماء هذا المحك إلى حد ما في حساب  
صدق اختبارات النكاه العام والشخصية وإلى حد كبير في حساب صدق  
اختبارات الاستعدادات المارقة والقدرات الطاقية. ومن المؤكد أن "الأعمال"

تختلف في مستوياتها وفي أنواعها. ويصدق هذا على الأصل الصناعية والتجارية والمهنية والعسكرية وغيرها، ورغم أن معظم مقاييس أداء " العمل " قد لا تعتبر محركات نهائية إلا أنها أفضل بكثير من معك سجلات التدريب الذي تحدثنا عنه آنفاً. ومع ذلك فقد لا يسمح قياس أداء العمل بتوحيد وتقنين الشروط التي تتوافر أثناء التدريب، كما أنه يتطلب تتبعاً طويلاً للمدى مما يؤدي إلى قلة عدد المفحوصين، وبالتالي يؤثر في انتنتج على نحو لا حيلة للباحث فيه.

### ج- الصدق التلازمي (أو المصاحب):

الصدق التلازمي أو المصاحب هو نوع الصدق الذي يدل على وجود علاقة بين درجات الاختبار ومؤشرات للمعك التي نحصل عليها في نفس الوقت تقريباً.

ولمّا لباحثون إلى الصدق التلازمي في بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبئي وما تتطلبه من وقت طويل، وفيها نعطي الاختبارات لمجموعة من الأفراد تتوافر لدينا عنهم معلومات خاصة بمعك معين، فنقارن مثلاً درجات طلاب الجامعة في اختبار معين بمتوسط تراكمي لدرجاتهم الدراسية حتى وقت الاختبار كما نقارن درجات الفصل في الاختبار بنجاحهم الراهن في العمل مثلاً.

ولمّا لباحثون إلى الصدق التلازمي في بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبئي وما تتطلبه من وقت طويل، وفيها نعطي الاختبارات لمجموعة من الأفراد تتوافر لدينا عنهم معلومات خاصة بمعك معين، فنقارن مثلاً درجات طلاب الجامعة في اختبار معين بمتوسط تراكمي

لدرجاتهم لدراسة حتى وقت الاختبار كما تقارن درجات العمال في الاختبار بلجاعتهم للراهن في العمل مثلاً.

وتتمثل أهم مشكلات الصدق التلزامي في صعوبة تحديد علاقات السبب والأثر، لأنه بدل بالفعل على علاقة للقران أكثر من دلالة على علاقة ' علوية '، ولذلك فإن استخدامه كبديل للصدق التلزمي يتضمن صعوبات منطقية ظاهرة، لأن التلزمي يتضمن في جوهره هذه العلاقة السببية.

ومن ناحية أخرى فإن للصدق التلزامي استخداماته الخاصة في أغراض معينة فهو أكثر ملائمة للاختبارات الشخصية والتي تلجأ إليها في حل المشكلات للراهن أكثر من مشكلات المستقبل. فعد إعداد مقاييس للشخصية بهتم الصدق التلزامي بمواضع مثل: هل هذا الطفل عصبي ؟، أما الصدق التلزمي فيجب على سवाल من نوع آخر وهو: هل من المحتمل أ، يصير هذا الطفل عصبياً ؟ ومع ذلك فكثر ما تنثار للتساؤلات حول وظيفة الاختبار في حالة الصدق التلزامي الذي تتصاحب فيه مطومات الاختبار ومطومات المحك والواقع أن الاختبار في مثل هذه الحالة عادة ما يكون بديل أبسط وأسرع وأقل تكلفة وجهداً من مطومات المحك. فإذا كان المحك يتطلب ملاحظة مباشرة لمرضى مدة أسبوعين من لحظة دخوله المستشفى، فإن الاختبار الصلاني صدقاً تلزامياً يمكنه أن يميز بين الأسوياء والعصابيين والحالات المشكوك فيها في وقت قصير وبذلك يختصر لنا صد هذه الأشخاص الذي يتطلبون هذه الملاحظة الشاملة، بالإضافة إلى اقتصاد الوقت والجهد الذي يبذل في الملاحظة المباشرة.

وقد يتضمن دليل الاختبار مطومات عن الصدق التلزامي إما على أنه بديل للصدق التلزمي (رغم صعوبات ذلك) أو كمؤشر للقوة الشخصية للاختبار.

وهذا للذرع من الصدق أكثر ملاءمة دون شك لاختبارات الشخصية ومع ذلك فقد يستخدم في حساب صدق اختبارات الاستعدادات والقدرات.

## 1- المجموعات المتضادة:

ومن للمحكات الأساسية التي نلجأ إليها في حساب الصدق للتأزمي ما يسمى " المجموعات المتضادة " وهو من أنواع المحكات المركبة التي تدل على الأكثر المتراكمة وغير المضبوطة للحياة اليومية، أي أنه لكل وضوحاً وأكثر تحيداً من المحكات الأخرى التي سبق مناقشتها، ويعتمد في أساسه على بقاء الفرد واستمراره داخل مجموعة سجن أو عزم بقاءه فيها وخرجه منها. ففي حسابنا لصدق اختبار الذكاء يمكن أن نقارن درجات ضعفات العقول المقيمين بإحدى المؤسسات بالدرجات التي يحصل عليها تلاميذ مدرسة عادية في نفس السن تقريباً. في مثل هذه الحالة يتأثر المحك بالعوامل العديد التي تتحكم في هذه المؤسسات بوجه عام. وكذلك الشأن إذا أردنا حساب صدق اختبار الاستعداد الموسيقي أو الاستعداد للميكانيكي فيمكن أن نقارن بين الدرجات التي يحصل عليها الطلاب الذين يدرسون الموسيقى أو الهندسة في مدرستين أو معنيين متخصصين، وبين درجات طلاب آخرين يدرسون في مدرسة ثانوية عادية أو كلية جامعية لم يتم اختبارهم على أساس الوظيفة السلوكية موضوع القياس.

وتستخدم طريقة المجموعات المتضادة في حساب صدق اختبارات الشخصية على وجه الخصوص، مثلاً يمكن أن نقارن أداء من يمارسون أعمالاً قبلية أو إقناعية بأداء من يمارسون أعمالاً كتابية في اختبار معين للمسات الاجتماعية ليمكن حساب صدق هذا الاختبار، مفترضين أن الأفراد الذين يصلون بهم من جهة تتطلب للعمل الاجتماعي (البيع، الكندروس... إلخ) وظلوا

يقومون بها سوف يتقنمون على غيرهم من الأكراد الذين يعملون في ميادين  
 أخرى لا تتطلب مثل هذا النوع من التعامل كالأعمال الكتابية والهندسية. بل  
 يمكن أن نقارن بين أداء الطلاب الذين يقومون بنشاط كبير خارج المنهج وأداء  
 غيرهم من الطلاب الذين لا يقومون بهذه الأنشطة خلال فترة دراسية معينة. كما  
 أننا كثيراً ما نلجأ إلى الجماعات المهنية لاصحاب صدق اختبارات الميول (ومنها  
 الميول المهنية) ومقاييس الاتجاهات. ومن الجماعات التي تستخدم في حساب  
 صدق مقاييس الاتجاهات الجماعات السياسية أو الدينية أو غيرها من الجماعات  
 الخاصة التي تمثل وجهات نظر مختلفة اعتقاداً ووضوحاً حول موضوعات  
 اجتماعية معينة. وقد تستخدم جماعات معينة مثل الأحداث المنحرفين المدعنين  
 في المؤسسات في مقابل الأحداث غير المنحرفين، أو المصابين في مقابل  
 الأصوياء، عند حساب صدق اختبارات الشخصية التي تتعلق بالتكيف الانفعالي  
 أو الاجتماعي. كما أن في بعض هذه الاختبارات يلجأ الباحثون إلى التشخيص  
 الطبي النفسي باعتباره أساساً لاختبار الأسئلة ودليلاً على صدق الاختبار، إلا أن  
 التشخيص الطبي النفسي لا بد أن تتوفر فيه عدة شروط هامة ليصبح محكاً  
 مفيداً، من أهمها أن يستند هذا التشخيص إلى ملاحظة طويلة المدى ودراسة  
 مفصلة لتاريخ الحالة، أما إذا اعتمدنا على مجرد مقابلة شخصية قصيرة أو على  
 فحص طبي نفسي غير عميق فإن يكون هناك سبب يدعو إلى توقع أن يكون  
 تشخيصنا أفضل من الدرجة التي يحصل عليها الفرد في الاختبار. نستدل من  
 أي منهما على الحالة الانفعالية للعميل. بعبارة أخرى يصبح التشخيص الطبي  
 النفسي كالاختبار مجرد مؤشر أو مقياس يحتاج في ذاته إلى حساب صدقه ولا  
 يكون محكاً نعتمد عليه في تقدير صدق الاختبارات الأخرى.



## 2- التقديرات:

وقد دلجا في حساب الصندق للتلامي لاختبارات الشخصية إلى التقديرات Ratings. وقد سبق أن تحدثنا عن التقديرات التي يعطيها المعلمون لتلاميذهم أو المشرفون على العمل لمعاملهم وغير ذلك عند حديثنا عن المحكات الأخرى والتي يطلبها بهدف الحصول على معلومات تتصل بهذه المحكات مثل التحصيل المدرسي الأكاديمي أو الأداء في نوع متخصص من الأعمال أو النجاح المهني عند حساب التنبؤي. إلا أن التقديرات قد تصبح في ذاتها محكاً تعتمد عليه في حساب الصندق للتلامي وخاصة إذا لم تقتصر على جانب معينة كالتحصيل مثلاً ولكن تتضمن حكماً شخصياً من جانب الملاحظ (أو الفاحص) على إحدى الصفات التي ترمع الاختبارات النقصية قياسها. ولذلك فقد نطلب من الفاحصين أن يعطوا تقديرات للمفوضين في خصائص معينة مثل السيطرة أو القيادة أو الأصالة أو الابتكار أو الأمانة... إلخ. وهذه التقديرات التي في ذاتها تعد محكات استخدمها الباحثون في حسابهم لصندق كثير من الاختبارات، وبخاصة لاختبارات الشخصية كما سبق أن قلنا، وذلك لصعوبة الحصول على محكات موضوعية في هذا الميدان وبخاصة بالنسبة للسمات الاجتماعية التي لا يمكن للحكم عليها إلا عن طريق التقديرات المستندة على العلاقة الشخصية.

ولواقع أن التقديرات التي نحصل عليها من أشخاص لديهم للتدريب والخبرة وفي ظروف مضبوطة ضيقاً تجريبياً دقيقاً توافر لنا معلومات عامة عن المحك ومما لا شك فيه أن زيادة ثقتنا في هذه التقديرات تتطلب الحصول على تقديرات مستقلة من أكثر من فاحص حتى يمكن التظلم على العوامل الذاتية التي تكرر في ثبات هذه التقديرات والتي سبق لنا مناقشتها في حديثنا عن ثبات

الاختبارات. كما يمكن أن تردّد دقة التقديرات إذا استخدمنا مقاييس التقدير rating scales بحيث تتضمن وحدات محددة تحديداً واضحاً مع وجود ضوابط خاصة للتخلص من أخطاء التقدير الشائعة مثل أثر أو وقع الهالة Halo Effect (وهي عبارة عن ميل للمدّرين أو الفاحصين إلى التّأثر بسمة واحدة مطلوبة، وتكون في صالح المفحوص أو في غير صالحه والتي تؤدي بالتّالي إلى التّأثير في أحكامهم الخاصة بالسّمات الأخرى)، كما لا بد أن تتوفر للمقيّرين " لفية بالسّمات موضوع التّقياس " بحيث يمكنهم الحكم على وجودها بدرجة معينة لدى الأفراد الذين نختبرهم. بعبارة أخرى لا يكفي أن يعرف الفاحص المفحوص لفترة طويلة ليصبح قادراً على إصدار حكم عليه وإنما يجب أن تكون قد نهيت له فرصة ملاحظته في مواقف مختلفة تظهر فيها السمة السلوكية موضوع الاهتمام. ولا يجب أن يطلى الفاحصون تقديرات عن المفحوص في سمة ليس لديهم عنها معرفة كافية.

وتلعب التقديرات دوراً هاماً في مقاييس المحك الذي يعتمد على التّقيوم الإكلينيكي. وفي مثل هذه الأحوال لا بد من تدعيم الحكم الإكلينيكي بعض المعلومات الإضافية مثل التاريخ المفصل للحالة ودرجات الاختبارات وغيرها من الوسائل الموضوعية، أما التّقيوم النهائي فيعتمد على عملية الحكم كما هو الحال في جميع التقديرات.

### 3- معاملات الارتباط بين الاختبار وغيره من الاختبارات:

ومن الوسائل الأخرى التي يستخدمها الباحثون في تقدير الصدق للتّلازمي حساب معاملات الارتباط بين درجات الاختبار التّجريبي (الجديد) وبين درجات اختبارات أخرى قائمة بالفعل. وقد تلجأ إلى هذه الطريقة إن كان الاختبار الجديد

صورة مختصرة أو مبسطة لاختبار موجود فعلاً وفي هذه الحالة يكون الاختبار الأصلي مقايماً للمحك. كما يمكننا أن نقرر صدق اختبار لفظي (ورقة وقلم) باستخدام اختبار ينوي علي لا يستغرق وقتاً طويلاً وله معامل صدق مرتفع ثم تقديره من قبل. وكذلك يمكن تقدير صدق اختبار جماعي باستخدام اختبار فردي (مثل اختبار ستانفورد بينيه الذي استخدم كثيراً كحكم لصاب صدق الاختبارات الجماعية للنكاه). وفي مثل هذه الأحوال يمكن اعتبار الاختبار الجديد أحسن الحالات صورة تقريبية للاختبار الأصلي. ويجب أن نلاحظ أنه ما لم يكن الاختبار الجديد بديلاً مبسطاً أو مختصراً للاختبار الأصلي لا يمكن أن نقوم باستخدام هذه الطريقة وإلا كانت نتائجنا موضع شك شديد.

#### صدق التكوين الفرضي:

يصدق بقصد التكوين للفرضي مدى قبيل الاختبار لتكوين فرضي معين أو سمة معينة. ومن أمثلة هذه التكوينات الفرضية النكاه والقيم الميكانيكية والطلاقة وسرعة المشي والخصابية والتلق... إلخ. ويعتمد هذا النوع من الصدق على وصف أوسع ويتطلب معلومات أكثر حول السمة السلوكية موضوع للقياس نحصل عليها من مصادر مختلفة.

وهناك أساليب كثيرة يمكن أن تلجأ إليها للحصول على أكبر قدر من المعلومات التي تلتقي الأعضاء على طبيعة السمة التي نهتم بها. ومن هذه الطرق ما يأتي:

## 1- تمايز العمر:

يمكن أن نلجأ إلى هذه الطريقة في حساب صدق اختبارات الذكاء (كاختبار ستيفرد بينيه، أو معظم اختبارات الذكاء لأطفال ما قبل المدرسة)، وذلك لتحديد ما إذا كانت الدرجات تتزايد بنظم العمر مفترضين أن القدرات تنمو بتقدير العمر خلال الطفولة وبالتالي فتوقع زيادة في درجات الاختبار إذا كان الاختبار صادقاً. ويعود الفضل إلى بيبيه في صياغة هذا المفهوم الخاص بتحديد للقياس العمري للذكاء.

ومن الواضح أن محك تمايز العمر لا يمكن استخدامه في الوظائف السلوكية التي لا تظهر تغيرات مطردة مع تقدم العمر ومن ذلك مقاييس الشخصية بوجه عام. كما أننا يجب أن نشير إلى أن هذا المحك محك ملبى وليس محاً إيجابياً. بمقارنة أخرى إذا لم نستطع درجات الاختبار أن تتزايد مع تزايد العمر فإن ذلك قد يدل على أن الاختبار غير صادق في قياسه للقدرات التي يدعي أنه صالحة لها، إلا أن تحسين الاختبار بحيث يصبح مقياساً متزايداً مع تزايد العمر لا يحدد لنا في ذاته الميدان الذي يقيسه الاختبار. فمثلاً يظهر كل من الطول والوزن زيادة منتظمة مع تزايد العمر ومع ذلك لا يمكن أن نحسب أحدهما أو كليهما من مقاييس الذكاء. وكذلك يجب أن نؤكد أنه في الاختبار النفسي الذي يقيس خصائص سلوكية تتزايد مع العمر إنما تتأثر هذه الزيادة أو النمو بظروف البيئة التي يتدن فيها الاختبار، فالظروف القبلية والحضارية قد تستثير بعض الخصائص السلوكية وقد تعطل البعض الآخر، ومعنى ذلك أن محك تمايز العمر لا يمكن أن نعتبره محكاً شمولياً بحال من الأحوال، وإنما هو كغيرها من المحكات محكوم بالموقف الحضاري الخاص الذي نشته منه.

## 2- معاملات الارتباط بالاختبارات الأخرى:

وقد دللنا في تقدير صدق التكوين الفرضي إلى حساب معاملات الارتباط بين الاختبار التجريبي وبين الاختبارات التي تتشابه معه في الوظائف التي يقومها والتي ثبتت صحتها، ويختلف معك حساب معاملات ارتباط الاختبار بالاختبارات الأخرى في صدق التكوين الفرضي عنه في الصدق التلازمي في أن معاملات الارتباط في صدق التكوين الفرضي يجب ألا تكون مرتفعة ارتفاعاً كبيراً جداً كما هو الحال في الصدق التلازمي، لأن الاختبار التجريبي الذي يرتبط ارتباطاً عالياً جداً باختبار جاهز دون أن تتوفر فيه أية مميزات إضافية يعني أن الاختبار التجريبي الجديد ما هو إلا ازدواج وتكرار للاختبار السابق لا مبرر له.

ويمكن استخدام معك معاملات الارتباط بالاختبارات الأخرى بطريقة أخرى لتحديد بها مدى ثمر الاختبار التجريبي من أثر العوامل التي لا علاقة بالوظيفة السلوكية موضوع القياس، أمثالاً إذا حصلنا على معامل ارتباط صفري بين اختبار معين للاستعدادات أو السمات الشخصية، وبين اختبارات الذكاء العام أو الاستعداد المدرسي واختبارات الفهم اللغوي فإن ذلك قد يعني أن هذه العوامل لا ترتبط بالاستعداد أو لسمعة موضوع القياس. ولذلك فكثر ما تتضمن كتيبات الاختبارات ومعاملات الارتباط باختبارات الذكاء العام والقراءة والفهم اللغوي على أنها مؤشرات غير مباشرة أو سلبية للصدق. ومن ذلك أنه إذا كانت هذه المعاملات مرتفعة فيكون ذلك مدعاة للشك في الاختبار، إلا أن معاملات الارتباط المنخفضة لا تعني في ذاتها أن الاختبار التجريبي صادق. ومهما يكن

الأمر فإن محك استخدام معاملات الارتباط في تقدير صدق التكوين الفرضي أقرب إلى طريقة تقدير الصدق التلازمي التي سبق الحديث عنها آنفاً.

### 3- التحليل العاملي:

وربما يكون التحليل العاملي من أهم الوسائل التي تستخدمها في تقدير صدق التكوين الفرضي. والتحليل العاملي (4) هو أسلوب إحصائي لتحديد السمات النفسية، وهو في أساسه يتضمن تحليل العلاقات بين الابعاد كما تشمل هذه العلاقة في صورة معاملات الارتباط نقطة البدء في التحليل العاملي حساب معاملات الارتباط بين كل اختبار والاختبارات الأخرى. وعندما نمحص الجدول الذي يتضمن هذه المعاملات (والذي يسمى هنا مصفوفة الارتباط) قد نلاحظ وجود تجمعات معينة بين الاختبارات مما يوحي بوجود سمة مشتركة بينها. فمثلاً إذا كانت الاختبارات التي ترتبط بعضها ببعض ارتباطاً عالياً من نوع اختبارات المفردات والمثابهاة والأضداد وتكملة الجمل وكانت معاملات ارتباط هذه الاختبارات بالاختبارات الأخرى منخفضة نستنتج من ذلك مجانباً وجود عامل النهم اللغوي. إلا أن هذه التحليل الذي يعتمد على مجرد فحص مصفوفة الارتباط عملية صعبة ومعقدة وغير مؤكدة بل معشوية وخاصة إذا كانت بطارية الاختبارات كبيرة، ولذلك نستخدم الأساليب الإحصائية الأكثر دقة لتحديد العوامل المشتركة التي يمكن اعتبارها مسؤولة عن معاملات الارتباط الناتجة. ولا مجال للإكراه في شرح هذه الطريقة الإحصائية الرفيعة في هذا السياق.

وفي التحليل العاملي نلاحظ أن عدد المتغيرات التي نصف في ضوءها أداء كل فرد تختصر باستمرار من عدد كبير من الاختبارات الأصلية إلى عدد

أصغر نسبياً من العوامل أو السمات المشتركة. ثم يمكن وصف للفرد في ضوء درجاته في العوامل القليلة المحدد. بعبارة أخرى يمكن القول أن أهم أعراض التحليل للعامل تبسيط وصف السلوك باختراق عدد المتغيرات، أي أنه يسير في اتجاه العلم الصحيح من التعدد والكثرة التي تتمثل في متغيرات الاختبارات إلى قلة التي تتمثل في العوامل المشتركة أو السمات.

#### 4- صدق المفردات:

وهناك طريقة أخرى لتقدير صدق التكوين لفرضي تتمثل في حساب صدق المفردات التي يتألف منها المقول. وقد دلجأ إلى هذه للطريقة على وجه الخصوص عند تقدير صدق اختبارات الشخصية. ويتلخص هذه الطريقة في اعتبار الدرجة الكلية في الاختبار نفسه محك الصدق. وقد نستعمل نوعاً من طريقة المجموعات المتضادة تتمثل في انتقاء جماعات متطرفة على أساس الدرجات الكلية في الاختبار نفسه محك الصدق. وقد نستعمل نوعاً من طريقة المجموعات المتضادة تتمثل في انتقاء جماعات متطرفة على أساس الدرجات الكلية في الاختبار ثم نقارن أداء المجموعة العليا بأداء المجموعة الدنيا في كل سؤال من أسئلة الاختبار. فإذا فشل السؤال في إعطاء نسبة " نجاح " من أفراد المجموعة العليا أعلى من أفراد المجموعة الدنيا اعتبره سؤالاً " غير صادق " وبالتالي يجب حذفه أو تعديله. وقد دلجأ إلى أسلوب معاملات الارتباط لهذا الغرض ذاته والذي يمثل في حساب معامل الارتباط الرباعي بين " النجاح والفشل " في كل سؤال من ناحية وبين الدرجة الكلية في الاختبار من ناحية أخرى، ولا يمكن أن نحفظ في الاختبار إلا بالأمثلة التي تعطي معاملات ارتباط دالة بين السؤال والاختبار. وإذا تبطلت هذه الطريقة في إعداد أسئلة

الاختبار فإن الاختبار ككل يمكن وصفه بأنه على درجة عالية من التماسق الداخلي حيث يميز السؤال في نفس اتجاه تمييز الاختبار. وقد دلجأ إلى تحويل هذه الطريقة بحساب معاملات الارتباط بين درجات الاختبارات الفرعية وبين الدرجة الكلية في الاختبار. لمعظم اختبارات الذكاء مثلاً تتكون من اختبارات فرعية عديدة تطبق منفصلة ثم نحصل على الدرجة الكلية للاختبار. وفي مثل هذه الاختبارات لا بد أن ترتبط درجات الاختبارات للفرعية بالدرجة الكلية، وتدل هذه المعاملات على التماسق الداخلي في الاختبار ككل.

ومن الواضح أن معاملات التماسق الداخلي هذه سواء اعتدت على أسئلة الاختبار أو على الاختبارات الفرعية إنما هي مفليسة للتجاسس بالمعنى الذي أشرنا إليه. ومن المؤكد أن درجة تجاسس أو تماثل الاختبار تتشابه إلى حد ما مع صدق للتكوين الفرضي لأنها تساعد نون شك في تحديد ميدان القصور أو القسمة التي يقسمها الاختبار والتي يعد عينة لها. ومع ذلك فإن التماسق الداخلي في حد ذاته ليس دليلاً كلفياً على الصدق لأنه إذا لم تتوفر لدينا وقائع خارجية ومستقلة عن الاختبار نفسه لا يمكن أن نعرف كثيراً على ما يقومه الاختبار.

### 5- الإجراءات التجريبية:

ومن الوسائل الأخرى التي يمكن أن دلجأ إليها لتقدير صدق التكوين الفرضي ما تطليه لنا الإجراءات التجريبية من معلومات حول أثر متغيرات معينة على درجات الاختبار. فإذا كان الاختبار التجريبي (الصنيد) معداً لقياس القابلية للقلق مثلاً يمكن أن نطبق هذا الاختبار على مجموعة من الأفراد نوضع في موقف يشهر القلق (مثل الامتحان في ظروف ضاحطة) ثم نحسب معامل الارتباط بين درجات الأفراد في اختبار القلق قبل التجريب وبين مختلف



للمؤشرات النفسية وغيرها التي تدل على القلق خلال التجريب (موقف الامتحان) وبعده. ويمكن القول أن هذا الإجراء التجريبي يشبه إلى حد كبير ما نلجأ إليه في تقرير المصنق للتبني، باستثناء الدور الإيجابي للباحث في تقديم مثير القلق والتحكم فيه. أما المقاييس المباشرة للقلق على القلق فيمكن اعتبارها محكاً تراجع عليه تذبذبتا من درجات الاختبار.

### موضوعية التصحيح:

بعد تصحيح الاختبارات بإعطاء الفرد درجة أو تقدير، وتفسير هذه الدرجة هو خطوة هامة، ولو أنه في ذاته يعد مقدمة لاتخاذ قرارا عملي أو تفسير علمي عن التردد أو مجموعة الأفكار موضوع القياس. وعلى كل حال فإن مشكلات تصحيح الاختبارات لا تقل أهمية وخطورة عن المشكلات الأخرى التي قرناها في هذا الكتاب، وإن كانت أكثر وضوحاً في نمط الاختبارات الذي نسميه "امتحانات العقل". ففي هذا النوع كثيراً ما يتعامل الشخص لماذا حصل على درجة منخفضة. ومن المؤكد أن السبب الرئيسي هو أنه في هذا النوع من الامتحانات يكون من الصعب تحديد الإجابة الجيدة وبالتالي يكون من الصعب إعطاء درجة مناسبة لاستجابة ليست صحيحة صحتاً كاملة.

ولا يمكن للباحث أن يجري دراسة السلوك كما لا يمكن أن يتخذ قراراً عملياً إذا لم تكن معايير التصحيح على درجة من الموضوعية، ولذلك يقترح البعض لطرح هذه المشكلة إما تحديد قواعد للحكم يلتزم بها المصححون أو استخدام الأسئلة للموضوعية التي تتطلب التعرف على الإجابة الصحيحة من بين عدة إجابات.

## تصحيح الأسئلة التي تتطلب استجابات حرة:

لا يزال القياس القوي يستخدم أسئلة ومشكلات تتطلب درجة من الحكم أثناء التصحيح، ومع ذلك فمن الممكن إعداد بعض الوسائل التي تتوفر فيها درجة من الموضوعية في هذا التصحيح وبخاصة فيما يتعلق بالخصائص الرئيسية للسلوك. فقد تقوم سلوك معين مثل الخطأ أمكن للباحثين إعداد نموذج التصحيح يحتوي على عينات من الخطوط تمثل مستويات مختلفة من الجودة وعلى المدرس أن يحدد العينة الأكثر شبهاً بخط التلميذ حتى يحدد الدرجة التي يعطيها له. وكذلك يمكن إعداد مقاييس مؤسسة على تقدير الإنتاج للحكم على نوع بعض ألوان السلوك الصناعي أو التجاري أو الإداري. ومع النجاح النسبي للوسائل الموضوعية في تصحيح بعض جوانب السلوك في الاختبارات التي تعتري على استجابات حرة، إلا أن هذه الوسائل لم تتجح نجاحاً مؤكداً في تصحيح الاختبارات اللغوية، ومع ذلك فإن التباين بين المصححين يمكن تقليله إذا أخذنا نماذج توضح نمط التصحيح المتفق عليه بالنسبة لإجابات نموذجية. ومن أمثلة ذلك نماذج التصحيح الخاصة بلغة اختبار ستفرد بينيه للذكاء، واختبار رورشاخ للشخصية واختبارات الابتكار الحديثة.

وعلى الرغم من أن نماذج التصحيح قد تكون مناسبة لأغراض القياس، إلا أننا نؤكد بعض جوانب الحذر حول تصحيح اختبارات الاستجابات الحرة بوجه عام. ومعظم مصادر الخطأ تنشأ في الواقع مما يمكن تسميته "المعلنة الشخصية" للمختبرين أو المصححين، فقد لوحظ مثلاً أن بعضهم يعطي درجات مرتفعة بصفة مستمرة بينما يميل البعض الآخر إلى إعطاء درجات منخفضة. كما أن الدرجة التي يعطيها المختبر أو المصحح تتأثر باستعدادة العقلي وتحيزه

وتتضمنه وأغراضه الشخصية وأهوائه، بل تتكرر بلثر الهالة التي تحيط بالمفحوص.

ويمكن التكامل من أثر هذه العوامل الذاتية إذا لجأنا إلى ما يسمى بالتصحيح ' الأصح ' أي يقوم المختبر أو المصحح بتقدير الدرجات للأفراد دون أن يعلم نتائج الاختبارات السابقة التي أجريت عليهم أو بالقلة التي ينتمي إليها في البحث (المجموعة التجريبية أو الضابطة) وبدون هذا الاحتياط يصعب علينا أن نتقارن تأثير العوامل الذاتية في الدرجة المعطاة.

**تصحيح الأسئلة التي تتطلب الاختبار من متعدد:**

يتكون نموذج التصحيح لما نسميه الاختبارات التعرف من قائمة الإجابات المصححة أو مفتاح التصحيح يمكن أن يستخدمه أي شخص. وقد اخترعت عدة وسائل أكثر كفاءة في الحصول على الاستجابات وتصحيحها، ومنها النسخ للكرتونية، وإعداد أوراق مستقلة للإجابة. ولا شك أن استخدام أوراق مستقلة للإجابة يؤدي إلى اختصار التكاليف، لأن كتابة الاختبار يمكن استخدامه أكثر من مرة، كما أن الإجابات يمكن تصحيحها بسهولة إما بمفتاح مقبوس أو بواسطة الآلات الحديثة. أما في النسخ للكرتونية فين الوجه الخلفي لسفحة الأسئلة (أو ورقة الإجابة) تلتصق به وتغفيه ورقة كرتونية، وبالطبع يطبع مفتاح الإجابة على هذا الوجه الخلفي. ويختار المفحوص الإجابة ثم يضع العلامة التي تدل على هذا الاختبار على الوجه الظاهر لورقة الأسئلة (أو الإجابة)، وبالطبع فين هذه علامة تطبعها ورقة الكرتون على الوجه الخلفي للورقة، وعند التصحيح يقوم المختبر أو المصحح بكشف الوجه الخلفي الذي توجد فيه مربعات مطبوعة توضح بأن يجب أن تكون الإجابات المصححة (أي مفتاح الإجابة)، وبالطبع

يكون من أيسر الأمور جمع عدد العلامات الكربونية التي طبعت في المربعات الصحيحة.

وقد ازداد في الوقت الحاضر الاعتماد على الآلات الإلكترونية في تصحيح الاختبارات. وفي هذه الحالة يقوم المفحوص بتسويد مكان الإجابة بالقلم الرصاص، وتحتوي الآلة على أصابع مكمية توضح علامات للقلم للرصاص، لأن الجرافيت الموجود في هذه العلامات يوصل التيار الكهربائي. كما تحتوي الآلة على عداد يوضح العدد الكلي للعلامات للموضوعة في أماكنها الصحيحة، كما أن الآلة يمكنها أن تسجل عدد الأخطاء، والدرجة التي تعتمد على (ص - خ)، وغير ذلك من أمثلة الدرجات وبخاصة ما يتعلق بتصحيح أثر للتخمين. ويمكن لبعض هذه الآلات أن يصحح لمئات من أوراق الإجابة تصحيحاً دقيقاً في الساعة. وقد تزيد عدد الأوراق الذي يمكن تصحيحه باستخدام الآلات الإلكترونية الحديثة.

أشهر معادلة لتصحيح أثر التخمين هي 
$$\frac{خ}{د - ص - ن - 1}$$

حيث تدل (ص) على الإجابات الصحيحة.

(خ) على عدد الإجابات الخاطئة.

(ن) على عدد الاحتمالات الاختبارات للإجابة عن السؤال الواحد.

والصعوبة الرئيسية التي تولدها آلات تصحيح الاختبارات، والتي يهتم بها المختصون في تطبيق الاختبارات، هي أن هذه الآلات لا يمكنها أن تصحح تصحيحاً دقيقاً ما لم تكن الأماكن المخصصة للإجابة قد قام المفحوص بتسويدها

تسريداً كاملاً ونظيماً وحسب المراسمات المطلوبة. فمن المعروف أن الألة تسطي وحدة الدرجة عندما تظهر العلامة التي وضعها المفحوص في المكان الصحيح ومع ذلك فنحن نتوقع أن تكون العلامة التي وضعها المفحوص بالقلم الخفيف أو الثقيل، أو تكون عريضة أو ضيقة. ولذلك فلا بد أن يضع المفحوص هذه العلامة بقلم معين وأن يملأ المسافة المخصصة للإجابة في ورقة الإجابة ملاً كاملاً حتى نتأكد أن الألة سوف تتعدها ولا تتجاهلها. بعبارة أخرى قد لا تتعامل الألة للعلامات غير الكاملة والمحو غير النظيف للأخطاء. وقد يكون عدد الإجابات في ورقة الإجابة التي لا تتوفر فيها المراسمات اللازمة كثيراً، ولذلك نجد في بعض المؤسسات موظفين يقومون بمحس كل ورقة قبل أن تسلم للألة لتصحيحها، وهؤلاء يقومون بمراجعة إجابات المفحوصين لا بفرض تصحيحها ولكن بفرض للتأكد من أن طريقة الإجابة مناسبة للتصحيح، ومعنى ذلك زيادة تكاليف التصحيح. ولا شك أنه يمكن التغلب على هذه الصعوبة بإجراء الاختيار وتطبيقه على نحو سليم.

#### المعليل

#### الدرجات الخام:

تصلبنا معظم الاختبارات وصفاً كميّاً مباشراً لأداء الشخص نطلق عليه اسم الدرجة الخام. وقد تكون الدرجة الخام عدد الأسئلة التي أجاب عليها المفحوص إجابة صحيحة أو الزمن الذي استغرقه في الأداء.

ويُفسر معظم الأفراد للدرجة الخام ويستخمنونها دون معرفة لجوانب  
لتصور فيها لأن هذه الدرجات تتميز بالسهولة في الحصول عليها وبمعرفة لها  
من خبراتنا الطويلة الامتحانات المدرسية.

وفي الواقع يمكن القول أن الدرجة الخام في الاختبار النفسي هي في حد  
ذاتها لا معنى لها وليس لها أية دلالة، ولا يمكن أن تُصر إلا بمقارنتها بمقياس  
معين. ومن المهم أن نؤكد في هذا الصدد أننا لا نستطيع تفسير الدرجات في  
الاختبار النفسي كما نفعل بالقيم التي نحصل عليها من مقياس للظواهر الطبيعية  
(الطول مثلاً). هـمقياس الظواهر الطبيعية لها صفر حقيقي ونحتوي على وحدات  
مساوية بتتابع المقياس، ولذلك فإن هذه للمقياس تسمح لنا بالمقارنة من  
الدرجات أو القيم التي نحصل عليها من المقياس مباشرة، فنقول مثلاً إن طول  
الطفل (ص) أصغر طول الطفل (ح)، أما في المقاييس للنفس لا نستطيع تقرير  
مثل هذا القول، فلنفرض مثلاً أن الطفل (ح) حصل على درجة مقدارها 10%  
من أسئلة اختبار في معاني الكلمات فهل يعني ذلك أن ذلك للطفل لا يعرف (لا  
عشر الكلمات التي يجب أن يعرفها ؟ الإجابة على ذلك بالنفي لأن الفاحصر قد لا  
يكون مهتماً إلا بالكلمات الصعبة غير المكلفة للطفل. وحتى لو حصل الطفل  
على الدرجة صفر في اختبار من هذا القبيل فإن ذلك لا يعني أن للطفل ليست  
لديه قدرة لغوية على الإطلاق. فالفرق بين الطفل (ح) الذي يحصل على  
الدرجة صفر في الاختبار والطفل (ص) الذي يحصل على النهاية القصوى، أي  
يجيب على أسئلة الاختبار جميعاً، هو فرق في القدرة على معرفة عدد محدود  
من الكلمات يتكون منها الاختبار من بين آلاف الكلمات.

معنى ذلك أن الفرق بين الأفراد في الدرجات الخام لا يدل بالضرورة على وجود مسافات \* حقيقية \* بين الأفراد. فمثلاً في اختبار الذكاء قد يحصل التلميذ (أ) على الدرجة 53 والتلميذ (ب) على الدرجة 56 والتلميذ (ج) على الدرجة 59، أي تكون الفروق في الدرجات الخام فروقاً مفسومة. فهل يعنى ذلك أن الفرق بين (ج) و (ب) يساوي تماماً الفرق بين (ب) و (أ) ؟ لا يمكن الإجابة على مثل هذا السؤال لإجابة مؤكدة، لأن الفروق بين الدرجات إنما تعتمد على الأسطة (أو الوحدات) التي نستخدمها.

والطريقة الوحيدة الممكنة التي يمكن بها أن نتحدث عن \* الفروق للمتساوية \* حديثاً له معنى تكون باستخدام محك عملي يحتوي على مستويات للقيم. ولتختلف المستويات يؤدي إلى قيم حدية مختلفة لنفس الاختبار، فالتلميذ الثلاثة السابق ذكرهم في مثالنا حصلوا على درجات خام تتساوى الفروق بينها في اختبار الذكاء، بينما نجد أن احتمال نجاحهم في مادة الهندسة مثلاً هو 0.70 ، 0.90 ، 0.96 على التوالي، بينما يكون احتمال نجاحهم في مهنة هندسية هو 0.0001 ، 0.50 ، 0.50 على التوالي. بحارة أخرى نصير \* للمساكنات المتساوية \* في أحد هذه المقاييس \* مسافات غير متساوية \* في مقياس آخر.

وعند تفسير درجات الاختبار يكون أمام الفاحص أن يختار أحد الاحتمالات الآتية:

- 1- أن يقارن الدرجة الخام مقارنة مباشرة بمستوى معين للذكاء، كما هو الحال في نظام للتهليات للصغرى في الامتحانات المدرسية.
- 2- أن يقارن الدرجة الخام بالدرجات الأخرى التي حصل عليها أفراد المجموعة التي اختبرها.

3- أن يقارن للدرجة الخام بدرجات مجموعة مرجعية تسمى عينة للتقنين بواسطة جدول المعايير Norms الذي يتم إصداره كخطوة هامة من خطوات بناء الاختبار.

4- أن يستخدم جدول التوقع لتقدير الأداء الممكن للفرد في المستقبل.

ومن بين هذه الطرق نجد أن الطريقة الثالثة أي طريقة مقارنة درجة الفرد بأداء عينة التقنين هي أكثرها شيوعاً. وعادة ما يرجع المصحح أو المحرر إلى جدول المعايير الخاص بالاختبار لمعرفة المدى العادي للأداء، ثم يحول الدرجة الخام إلى نوع معين من الدرجة المشتقة والتي تعتبر سجلأً دائماً للوضع النسبي للفرد، وأشهر أنواع الدرجات المشتقة هي السكيزات والدرجات المعيارية والأعمار العقلية ونسب الذكاء.

#### المنهيات

أبسط وسائل للمقارنة هي أن نرتب الدرجات التي يحصل عليها الأفراد إما ترتيباً تنازلياً أو تصاعدياً. وبذلك يمكن أن نحدد الوضع النسبي للفرد بالنسبة لأقرانه. ويعتمد الترتيب دون شك على عدد الأفراد في الجماعة، فإذا أردنا أن نحدد التمييز في وضع الفرد من منسبة معينة لأخرى تنشأ في هذه الحالة صعوبات خاصة باختلاف حجم المجموعات، ولتجنب هذه الصعوبات نحول الترتيب إلى درجات منيئية (وأحياناً تسمى رتب منيئية). وللدرجة المنهية هي رتبة صورة نسبة منوية. وللدرجة المنهية للشخص تدل على نسبة الأفراد من بين المجموعة التي تقع في مرتبة أقل منه. فإذا فترارنا أن لدينا 40 شخصاً منهم 27 شخصاً حصلوا على درجات أعلى من الشخص (أ) ر 12 شخصاً حصلوا على درجات أقل منه، فإننا في هذه الحالة نضع هذا الشخص (أ)



لو أي أشخاص آخرين في مرتبة قريبة منه بين المجموعتين معتبرين أن 27 حالة تقع أعلى منه و 12 حالة تقع أدنى منه. وحيث أن 12 هي 321% من الحد للكي البالغ 40 شخصاً فلنا في هذه الحالة نقول أن الدرجة المنبئية للشخص (1) هي 31.

وبطريقة الحساب هذه فإن الشخص الذي يقع في منتصف الجماعة تماماً هو الذي يحصل على العتني الخمسين، وهو ما يجر عنه إحصائياً باسم الوسيط، وهو في الواقع يدل على أداء الفرد النموذجي في الجماعة. وبمنس الطريقة يمكن للحدث عن مقاييس الترتيب النسبي المعادلة مثل الترتيبات أو الترتيبات وغيرها.

ومما لا شك فيه أن تحويل الدرجات الخام إلى درجات منبئية يغير من شكل التوزيع. فمن المعروف أن التوزيع التكراري للنموذجي للدرجات الخام هو المنحني الاعتدالي أو ما هو أقرب إليه. حيث يكون أكبر عدد من الحالات في المنتصف (أو حول المتوسط الحسابي) ويقل عدد الحالات تدريجياً كلما اتجهنا نحو طرفي التوزيع. أما إذا تحولت جميع الدرجات الخام إلى درجات منبئية فإن التوزيع التكراري يصبح أقرب إلى المنحني المستطيل. ومضى ذلك أن الحالات لكثيرة التي تقع في المنتصف تنتشر في حين أن الحالات القليلة التي تقع في الأطراف تتجمع بعضها مع بعض.

ولا بد أن يشير إلى أننا لا نستطيع استخراج متوسطات للدرجات المنبئية، لأن هذا المتوسط يختلف في مقداره عما لو استخرجنا المتوسط الحسابي للدرجات الخام ثم حولناه إلى درجة منبئية. وبدلاً من اللجوء إلى المتوسط الحسابي لتولس النزعة المركزية للدرجات المنبئية يمكن الباحث أن

يعتمد على الوسيط باعتباره المقياس الإحصائي المناسب للزعة المركزية في هذه الحالة.

كما يجب الإشارة إلى أن الدرجات المتبينة لا يمكن المقارنة بينها ما لم تكن الجماعات التي تشتق منها يمكن المقارنة بينها بالفعل. فإذا اختلفت الجماعات، كل تكون إحداها مثلاً تلاميذ مدرسة ثانوية صناعية والثانية صال تلمذة صناعية والثالثة طلاب كلية الهندسة، فقد تكون الدرجة الخام التي تقابل الدرجة المتبينة لـ 60 مثلاً لطلاب الجامعة تقابل مثلاً المتبيني 75 في مجموعة طلاب المدرسة الثانوية للصناعية وتقابل كذلك المتبيني 90 في مجموعة صال التلمذة الصناعية. بعبارة أخرى إذا لجأنا إلى استخدام الدرجات المتبينة لا بد للباحث أن يضع في اعتباره المجموعة المرجعية أو هيئة التقيين التي اشتمت منها هذه المعايير، ولو أن ذلك لا يحول دون استخدام مجموعة من جداول التحويل إلى درجات متبينة تسمح للمختبر أو الباحث أن يقارن الشخص الذي يختبره بمجموعات مرجعية مختلفة.

#### العمر العقلي ونسبة النكاح:

تعتمد بعض الاختبارات على مقياس العمر في تفسير درجاتها، ومن أشهر هذه الاختبارات ستانفورد بينيه، وفي هذه الحالة يتم تجميع أسئلة الاختبار حسب مستويات العمر، فمثلاً الأسئلة التي يجيب عليها أغلبية الأطفال من سن 7 سنوات، وذلك التي يجيب عليها أغلبية الأطفال من سن 8 سنوات تعتبر ملائمة لمستوى 8 سنوات وهكذا، ومعنى ذلك أن درجة الطفل في مثال هذا الاختبار تقابل أعلى مستوى عمري يمكنه أن يصل إليه بنجاح. فإذا استطاع طفل عمره 9 سنوات أن يجيب بنجاح على أسئلة صر 12 منه يكون صره العقلي 12

بصرف النظر عن عصره الزمني. وهو في هذه الحالة يكون سابقاً لعصره الزمني بثلاث سنوات ما دلم أدائه يكافئ أداء طفل متوسط من عصر زمني أكبر منه بثلاث سنوات.

وقد لوحظ أن أداء الفرد في الاختبارات التي تعتمد على مفهوم العمر يظهر مقدراً معيناً من التنشت. بعبارة أخرى قد يشغل الطفل في أداء بعض الأسئلة التي تقع في مستوى أدنى من عصره العقلي، كما يمكنه الإجابة على بعض الأسئلة التي تقع في مستوى أعلى منه. ولذلك نلجأ عادة إلى حساب العمر القاعدي أو الأساسي أي أعلى مستوى عصري يستطيع الطفل أن يجيب على جميع أسئلته والأسئلة السابقة عليه ثم تضاعف إليه درجات جزئية، بالطبع في صورة كسور من السنة أو الشهر، تدل على جميع الأسئلة التي لجاب عليها من مستويات أعلى من العمر القاعدي، وبذلك يصبح العمر العقلي للطفل هو مجموع العمر القاعدي والأشهر الإضافية التي تدل على ما اكتسبه الطفل من مستويات عصرية أعلى.

وقد نلجأ إلى استخدام معايير العمل العقلي في اختبارات لا تنقسم إلى مستويات حسب السنوات. وفي هذه الحالة نحصل أولاً على الدرجة الخام التي قد تكون العدد الكلي للأسئلة الصحيحة في الاختبار ككل أو قد تكون معتمدة على الزمن أو عدد الأخطاء أو مزيج منها. ثم نحسب متوسطات الدرجات الخام التي يحصل عليها الأطفال في كل مجموعة عصرية من بين عينة التفتين، وتكون تلك المتوسطات هي معايير العمر لذلك الاختبار.

ويجب أن نؤكد أن وحدة العمر العقلي لا تبقى ثابتة بتقدم العمر ولكنها تسهل إلى الضمور في السنوات المتأخرة عنها في السنوات المبكرة من النمو.

فلذا كان الطفل متخلفاً عاماً واحداً في سن 4 سنوات فإنه يصبح متخلفاً ثلاث سنوات في سن 12 سنة. فعمام ولحد من " النمو " العقلي في الفترة من سن 3 حتى 4، يعادل 3 سنوات من النمو في الفترة من سن 9 حتى سن 12، وحيث أن النمو العقلي يتقدم بسرعة في السنوات المبكرة من النمو ثم يتناقص تدريجياً بالتقارب الفرد من حد النضج، فإن العمر العقلي يضمن تبعاً للعمر. وهذه العلاقة يمكن فهمها بصورة أوضح إذا تصورنا طول الشخص على أنه يمكن صياغته في صورة " عمر طولي " نجد أن الفرق بالسنتيمترات بين " العمر الطولي " في عمر 3 أو 4 يكون أكبر من ذلك للفرق بين " العمر الطولي " في سن 10 و 11. ومعنى ذلك أن عاماً واحداً من التفوق أو التخلف في سن 5 سنوات يمثل انحرافاً عن المستوى المتوسط أكبر كثيراً من مقدار مماثل له من التفوق أو التخلف في سن 10 سنوات.

ومن أجل الحصول على مقاييس يسمح بتصوير موحد بصرف النظر عن العمر الزمني للمعرض لبتكر ترميز مفهوم نسبة الذكاء رغم أن الحاجة إلى مثل هذه النسبة أشار إليها قبله شترن وكولمان، كما ستوضح فيما بعد، ونسبة الذكاء هي نسبة العمر العقلي إلى العمر الزمني مضروبة في 100 للتخفيف من الكسور العشرية، على النحو الآتي:

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلي}}{\text{العمر الزمني}} \times 100$$

فلذا تعادل العمر العقلي والعمر الزمني تكون نسبة الذكاء في هذه الحالة 100 وتدل على أداء متوسط أو عادي، وإذا قلت عن 100 دلت على تخلف أو

ضعف، وإذا زادت عن 100 نلت على الترقى. وتعد نسبة الذكاء أيضاً في ضبط الضمور في وحدة العمر العقلي ضبطاً تلقائياً باستخدم النسبة. فمثلاً إذا كان لعمر العقلي لطفل عمره الزمني 4 سنوات هو 3 تصبح نسبة نكته في هذه الحالة 75 فإن هذا الطفل حينما يصل إلى سن 12 سنة يصبح عمره العقلي 9 وتظل نسبة نكته ثابتة أي 75. أي أن نسبة للذكاء تثل على نفس الوضع للنسبي للفرد في الجماعة سواء في سن 4 سنوات أو في سن 12 سنة. وبعبارة أخرى فإن نسبة الذكاء يمكن استعمالها في المقارنة في مختلف الأعمار، بمعنى أن تفسير نسبة الذكاء يظل هو نفسه بصرف النظر عن العمر الزمني للمفحوص.

ولا بد أن نشير في هذا الصدد إلى أن نسبة الذكاء تظل ثابتة في حالة واحدة فقط وهي أن وحدة العمر العقلي تتناسب تناسباً مباشراً مع العمر الزمني. وهذا شرط ضروري لتمكين استعمال نسب الذكاء استخداماً له معنى. ومعنى ذلك أنه حينما تتناسب وحدة العمر العقلي فإن الفروق الفردية التي تقاس بهذه الوحدة سوف تتزايد بنفس النسبة. ويشبه ذلك من الناحية الحسابية أن الفروق الفردية في الطول تكون أكبر 12 مرة حينما تقاس بقبوصات عنها لو قيمت بالأقدام وذلك لأن القبوصة هي  $\frac{1}{13}$  من القدم وبفرض المنطق فإن الفروق الفردية في العمر العقلي يجب أن تكون في سن 14 ضعفها سن 7 سنوات، ولا يمكن لنسب الذكاء أن تظل ثابتة إلا إذا تزايد الانحراف المعياري للأعمار العقلية بنفس نسبة زيادة الأعمار الزمنية، وبهذا الشرط وحده يمكن أن يكون لنسبة الذكاء نفس المعنى في مختلف الأعمار.

وقد أمكن تحقيق هذا الشرط في اختبار ستانفورد بينيه فقد اتضح من دراسة انتشار الأعمار العقلية من سن 6 حتى سن 18 أن هناك اتجاهًا نحو زيادة القياس في الأعمار العقلية تبعاً لزيادة الأعمار الزمنية، ولكننا نجد في كثير من الأحيان عدم تساوي الانحرافات المعيارية لنسب الذكاء في مختلف الأعمار، وهذه الصعوبة لم تكن لتغلب عليها في اختبار ستانفورد بينيه بإعداد جدول تصحيح يستخدم فيه تصحيح نسب الذكاء في مستويات عمرية معينة.

وهناك مشكلة أخرى في تطبيق معيار نسبة الذكاء نجدها في اختبار البالغين والكبار. فالواقع أن مفهوم معايير العمر يدل دلالة مباشرة على أن فاقته مقصورة على الأطفال، أما بالنسبة للكبار فإننا نجد أن في اختبار مثل ستانفورد بينيه لا يتمسك أداء البالغ المتوسط كثيراً أبعد من مستوى 15 سنة. وعلى ذلك فإن العمر العقلي للبالغ المتوسط هو حول هذا المقدار (حوالي 15 سنة و 9 أشهر). وقد اضيف إلى الاختبار مستويات ودرجة لذكاء الكبار حتى يتوفر في الاختبار سقف مناسب وينبغي يمكننا الحصول على أعمار عقلية أعلى من 15 سنة و 9 أشهر، ولكننا إذا وجدنا شخصاً معيناً يحصل على عمر عقلي مقداره 20 في هذا الاختبار فإن ذلك لا يسمح لنا بتفسير مؤكد وعلى نفس الدرجة من الوضوح كما نجد في حالة الأعمار العقلية 8 أو 10. فمن المؤكد أن للعمر العقلي 20 لا يمكن تحديده بما يمكن أن يؤديه لشخص المتوسط البالغ من العمر 20 سنة، لأن الشخص المتوسط للبالغ من العمر 20 سنة يحصل على عمر عقلي مقداره 15 سنة و 9 أشهر. أما إذا اختبارنا شخصاً بلغاً ضعيف العقل يقل عمره العقلي عن 15 سنة و 9 أشهر فإن مفهوم العمر العقلي لا يختلف في استخدامه في هذه الحالة عنه حينما نستخدمه مع الأطفال. أما في حالة البالغين الأسوياء أو المفقوتين فلا مناس من استخدام أنواع أخرى من الدرجات

مثل المقياسات (التي أشرنا إليها) أو للدرجات المعيارية (التي منشور إليها فيما بعد).

أما المشكلة الثالثة التي تواجهنا حينما نقرر الدرجات الخام بدرجات العمر تتمثل في أن هذا النوع من المعيار لا يستخدم إلا في الوظائف النفسية والمساكنات المعقدة التي تتغير تغيراً واسعاً ومتسقاً مع العمر والنمو. أما بالنسبة للمساكنات التي لها علاقة ضئيلة أو لا تكون لها علاقة بالعمر فلزمنا لا يجوز مطلقاً تحويل الدرجات الخام لمقاييسها إلى درجات أعمار. ومن أمثلة الاختبارات التي يستحيل تطبيق معايير الأعمار عليها الاختبارات التي تقيس سمات الشخصية وخصائصها.

#### الدرجات المعيارية:

تعتبر الدرجات المعيارية أفضل صورة تحويل للدرجات الخام. وهي تعبر عن بعد الدرجة التي يحصل عليها الفرد عن المتوسط في ضوء الانحراف المعياري للتوزيع. وللحصول على الدرجة المعيارية لا بد أولاً من حساب المتوسط الحسابي الذي هو الوسط الحسابي الذي يحصل عليه جميع الدرجات ثم قسمة هذا المجموع على عدد الأفراد. ثم حساب الانحراف المعياري (الذي هو مقياس للتشتت أو الانتشار) وهو في أصله متوسط الانحرافات درجات المفحوصين كأفراد عن المتوسط الحسابي للجماعة. وبذلك الانحراف المعياري على مقدار الاختلاف داخل الجماعة على عكس المتوسط الذي يدل على التفرقة المركزية للدرجات. وبعد حساب كل من المتوسط والانحراف المعياري يمكن إيجاد الفرق بين الدرجة التي يحصل عليها الفرد ومتوسط الجماعة ثم قسمة هذا الفرق على الانحراف المعياري للمجموعة، وبهذه الطريقة

نستبدل الدرجة الخام بدرجة مشتقة نقل على وضع الفرد بالنسبة لمتوسط الجماعة، وبفقدنا الانحراف المعياري في تحديد مقدار البعد عن هذا المتوسط ارتفاعاً أو انخفاضاً، وبالتالي نستطيع أن نقول أن الفرد (أ) أعلى من المتوسط بمقدار 205 وحدة انحراف معياري، وأن الفرد (ب) أقل من المتوسط بمقدار 1- وحدة انحراف معياري.

ومن المشكلات التي تعانيها الدرجات المعيارية أن بعض قيمها قد تكون سالبة كما أنها تتضمن كسوراً من الواحد الصحيح مما يسبب خلطاً وصعوبة في استخدامها سواء في الجانب الحسابي الخاص بحساب هذه الدرجات أو الجانب الخاص بتفسير هذه الدرجات والاستفادة بها عملياً. ولهذا السبب اقترحت عدة تحديلات من أجل وضع الدرجات المعيارية في صورة أكثر ملاءمة، منها مثلاً الدرجات الثنائية التي تقترض متوسطاً مقداره 50 وأن كل وحدة انحراف معياري تساوي 10. (ومن المعروف أن متوسط الدرجات المعيارية صفر وأن انحرافها المعياري هي واحد صحيح).





# اختبارات الذكاء العام



## اختبارات الذكاء العام

• توجد فئة من الاختبارات النفسية تسمى اختبارات الذكاء العام. وتتميز هذه الاختبارات بأنها تستخدم في مواقف متنوعة ويتحدد صدقها ليس ضوء محكات أكثر شمولاً. كما تتميز هذه الاختبارات بأنها تعطي درجة واحدة مثل نسبة الذكاء، لتقل على المستوى العقلي العلم للمفحوص، ويرى ترمان وميريل أن الوصول إلى هذا التقدير الشمولي للمستوى العقلي يتطلب "تصويب الرماح إلى الأهداف الحاسمة". وبعبارة أخرى يعرض على للمفحوص عدداً كبيراً ومتنوعاً من الأعمال على أساس أنها تمثل عينة ملائمة من أساليب الأداء العقلي. ومع ذلك فإن البحوث الحديثة تؤكد أن هذه الاختبارات تتسبب بوظائف عقلية معينة مثل القدرة اللغوية، وتتجاهل تماماً بعض الوظائف العقلية الأخرى.

• وكثيراً ما تسمى اختبارات الذكاء العام باختبارات الاستعداد المدرسي لأن صدقها يتحدد في العادة في ضوء محكات التحصيل الأكاديمي. كما أن هذه الاختبارات تستخدم في أغلب الأحوال في أغراض التصفية للمهنية العامة (سواء من الناحية القريبية أو المهنية أو العسكرية)، ولمفروض أن يتبعها تطبيق اختبارات الاستعدادات الأكثر تخصصاً. ويشيع هذا على وجه الخصوص في اختبار المراهقين والراشدين الأسوياء عند الإرشاد أو الانتقاء أو التصنيف أو التوجيه أو غير ذلك من الأغراض التي أثبتنا إليها في الفصل الأول من هذا الكتاب، كما يوجد استخدام آخر لاختبارات الذكاء في الأغراض الإكلينيكية خاصة وهو تحديد الضعف العقلي وتشخيصه. وفي مثل هذه الأحوال فلن استخدام الاختبارات القريبية مثل مقاييس ستانفورد بينيه ومقاييس وكسلر أكثر شيوعاً من غيرها.

وسوف نتناول في هذه الوحدة المقاييسين بشيء من التفصيل، دون أن نتجاهل الاختبارات الأخرى التي سمحت إلى أيرلس الفكاك العام.

### الفرد بينيه

الفرد بينيه هو العالم لهذا الأصل، بدأ تدريبه في ميدان الطب ثم سرعان ما أصبح أشهر علماء النفس في فرنسا في عصره، وله فصله العظيم في علم النفس في بلد. فقد أنشأ زميله بيوني أول معمل لعلم النفس عام 1889 في جامعة السوربون، وأول مجلة متخصصة باللغة الفرنسية 'حداية علم النفس عام 1895'.

وقد بدأ بينيه عالماً تجريبيًا، فقبل أن يعد اختباره العقلية المشهورة اهتم بدراسة العمليات العقلية للمروية، وأطلق على منهجه اسم 'سيكولوجية الفرد' وفي رأيه أن أهم ما نجده من فروق بين الناس إنما هو في الأنشطة العقلية العليا كالاستدلال والحكم، وهي جميعاً من المسائل التي يصعب قياسها بالوحدات الفيزيائية (كالإحساس) كما لا تتوافر فيها الوحدات الأولية للسلوك الذهني، والتي كان يهتم بها أغلب علماء النفس من أواخر ثمانية فوننت (وملهم تتشورا).

وقد تأكدت صحة وجهة نظر بينيه من نتائج بعض الدراسات التي أجريت على صدق اختبارات جيمس ماكين كاتل (التي اقتصرت فيها على المسير فرميس جالتون في إنجلترا) كمقاييس عقلية، وبخاصة بحوث كلارك وزلو فرميس جالتون في إنجلترا). ولذلك بدأ بينيه يطبق في بحثه عن الاختبارات العقلية ما توصل إليه من نتائج البحث التجريبي في صلابة التفكير، فانتقد

لختبارات كمال وجاقتون لأنها تقتصر على الذواحي الحسية البسيطة، والفضل الاختبارات الأكثر تعقداً وتركيباً، واقتراح عدداً من الاختبارات لقياس " الملكات العقلية ".

وقضى بينيه لفترة بين عامي 1896 و 1905 في إجراء سلسلة طريقة من البحوث على الأطفال الأسوياء وغيرهم مجرباً أنواعاً جديدة من الاختبارات لمختلف " الملكات "، بالإضافة إلى عدد من التجارب حول المسائل الأخلاقية في سيكولوجية التفكير في ذلك الوقت مثل التجريد، والتفكير بدون صور وزمن الرجوع والتمييز الحسي. وكان الهدف الذي يسعى إليه بينيه من ذلك هو تحديد العلاقة بين هذه الاختبارات والعمر الفعلي والتحصيل المدرسي وتغيرات الممطين للفترة العقلية. وكانت بحوثه جيدة التصميم وأدت به إلى التحول من الاهتمام بالاختبارات التي تقيس " الملكات " التقليدية إلى تلك التي تقيس المستوى العقلي بالفعل.

وهكذا كانت نظرة بينيه إلى الذكاء أكثر شمولاً، حين لجأ إلى استخدام درجة كلية واحدة (العمر العقلي) في قياس الذكاء.

نشأة مقياس بينيه:

في عام 1904 طلبت وزارة المعارف الفرنسية من ألفرد بينيه والطبيب الفرنسي ثيودور سيمون دراسة الطرق التي يمكن أن تستخدم في تصنيف الأطفال الذين يتأخرون ببطء التعلم في المدارس الفرنسية، فحكفا على وضع مقياسهما المشهور للذكاء بغرض التمييز بين الأطفال الأسوياء وضعفاء العقل.

ولم يكن الهدف دراسة وتحليل استجابات المتخلفين عقلياً، وإنما كان الهدف عملياً بحثاً وهو تقدير وقياس ذكاء هؤلاء الأطفال بصفة عامة وتحديد مستوياتهم العقلي ومقارنتهم بالأطفال الأسوياء من نفس العمر والمستوى. أي الهدف كما يقول بيديه هو بناء مقياس " متري " للذكاء يتكون من عدد من (الاختبارات) المترتبة في الصعوبة، مرتبة حسب الأسهل حتى يستطيع الفاحص أن يحدد مدى تقدم الطفل عقلياً أو تخلفه. وقد سعى بيديه إلى أن تكون هذه " الاختبارات " مقبوعة، مع التركيز على ذات معينة من العمليات العقلية اهتم بها دون غيرها عدد إعداد المقياس وهي: للحكم والفهم العام والمبادأة والقدرة على التكيف.

ويتكون مقياس بيديه الأصلي الذي نشر عام 1905 بأمر من نقولس بيتيغ - سيمون للذكاء من 30 لغزاً، وفي يلي قائمة بهذه الاختبارات، نذكرها كاملة لندين إلى أي حد اختلفت عن اهتمامات علماء القياس العقلي قبل بيديه، وكيف أنها هأت الموقف لظهور عدد كبير من الاختبارات العقلية تتحدر في معظمها " من صلب " مقياس بيديه على حد تعبير أمتنازي بالإضافة إلى شيوعها في معظم تحليلات مقياس بيديه ذاته.

- 1- القزور البصري (تتبع شمعة مضيئة بالرأس والعين).
- 2- الفهم عن طريق اللمس (الإصبع يمسك بمكبب بعد لعمه).
- 3 الفهم عن طريق البصر أو اللمسة.
- 4- التعرف على الطعام (التمييز بين قطعة الخشب وقطعة لشكولاته).
- 5- البحث عن الطعام (الاستجابة لقطعة من الشوكولاته ملفوفة في الورق).
- 6- تتبع التعليمات البسيطة أو الإيماءات المتكررة.

- 7- الإشارة إلى الأياء وخلصه أعضاء الجسم (الرأس، الألف ... إلخ).
- 8- التعرف على الأشياء من صورها.
- 9- تسمية الأشياء في الصورة.
- 10- التمييز الحسي بين طول خطين.
- 11- تكرار ثلاثة أرقام.
- 12- التمييز بين وزن شيئين.
- 13- القابلية للإيهام (المقارنة بين خطين متساويين، لو البحث عن أشياء لا وجود لها في صورة).
- 14- تعريف الكلمات البسيطة (التي تكل على أشياء مألوفة).
- 15- تكرار جملة تتكون من 15 كلمة.
- 16- بيان أوجه الاختلاف بين شيئين.
- 17- الذاكرة البصرية.
- 18- رسم الأشكال من الذاكرة.
- 19- مدى ذكر الأرقام.
- 20- بيان أوجه التشابه بين الأشياء.
- 21- التمييز بين الخطوط بسرعة.
- 22- ترتيب 5 أوزان.
- 23- تحديد الوزن الدقيق (من بين الأوزان الخمس في الاختبار السابق).
- 24- إعطاء كلمات ذات دلالة واحدة.
- 25- تكملة الجمل.
- 26- بناء جملة تحثري على 3 كلمات معينة.
- 27- الإجابة عن الأسئلة (مثل: ماذا تفعل وأنت تمشي؟).
- 28- تحديد الوقت بعد أن يمل عقارب الساعة محل بعضها.



29- نقي وتطبيع قطعة من الورق.

30- التمييز بين الكلمات المجردة (مثل: حزن وسلم).

ولذا ففحصنا هذه القائمة فحسباً دقيقاً نجد أن الاختبارات الثلاث الأولى هي في الواقع اختبارات للتصور الحركي، أما الاختبارات الأخرى فيمكن أن تعد من نوع اختبارات الأداء الأكسي. ومن بين هذه الاختبارات نجد 18 اختباراً (حرفي الفلشئين) تقيس ما يسمى بـ جيلفورد التفكير المعرفي، كما نجده اختبارات تقيس الذاكرة بالملامح المختلفة. ولا يوجد إلا 3 اختبارات فقط تقيس ما يسمى في الوقت الحاضر بـ تفكير الابتكاري (أو الإنتاج القاعدي عند جيلفورد).

وكانت الطريقة التي استخدمها بيليه في تقنين مقاييسه غير ملائمة، فقد طبقه على 50 طفلاً (10 من كل فئة من فئات الأعمار 3، 5، 7، 9، 11) اختارهم معطوهم على أساس أنهم متوسطو القدرة العقلية، بالإضافة إلى عدد لم يذكره بيليه من الأطفال ضعاف الحول. وكانت طريقة التصحيح تعتمد على أساس تجريبي مبني يسمح بتصنيف الأطفال تبعاً لأعلى مستوى يصل إليه الطفل، فالمعتوهون لا يمكنهم تجاوز الاختبار رقم 6. ولا يتجاوز البلهاء الاختبار رقم 15، أما حدود المألوفتين للممكن معروفة، رغم أنهم معجزون في العادة عن الإجابة على الأسئلة المجردة التي يستطيع معظم الأطفال للعالمين من سن 11 الإجابة عليها.

\* من المحتمل أن ننجح في يوم ما في إيجاد علامات على التنفص الميكولوجي مستقلة تماماً عن العمر. وقد يكون من المفيد جداً التعرف على هذه العلامات إلا أننا في الوقت الحاضر نجد أن ما يجذب للتباعد أكثر من غيره هو التشابه بين الأسوياء من صفراء السن وغير الأسوياء من كبار السن. وهذا

التشابه من الكثرة والغلبة بحيث أننا لو حصلنا على امتحانات طفل لا نعرف صرء لا نستطيع الحكم عليه بالسواء أو عدم السواء<sup>٢</sup>.

للتحليل السيكولوجي التقليدي مثل الملكات أو عقبات الإحساس أو غيرها. وبالإضافة إلى ذلك فإن وضع اختبار في المقياس أو حذفه أو تعديله كان من المسائل التجريبية، ويدل على خبرة بينيه الطويلة بالتجريب العملي. فقد كان يهتم بمدى العلاقة بين الاختبار والحكم المحك (أي بنوع من الصدق). ثم إلى هذه الاختبارات لم تكن محدلات معملية دقيقة وإنما كانت أدوات عملية بسيطة وهامة لدراسة الأطفال في مواقف واقعية (مثل حجرة الدراسة).

#### مقياس ستانفورد - بينيه لثلاثين

لم تكن محاولة جودارد ترجمة ونقل مقياس بينيه إلى اللغة الإنجليزية هي المحاولة الوحيدة، ففي عام 1910 نشر هوي ووربل ترجمة له، كما نشر فالفين وكولمان ترجمة أخرى في عام 1911. وفي عام 1913 بدأ بيرت نقل المقياس إلى الإنجليزية وتقيمه على عينة كبيرة من الأطفال الإنجليز عندما عين أخصائياً نفسياً لمقاطعة في لندن، وظل يتابع مشروعه هذا طوال تاريخه العلمي كأستاذ لعلم النفس بجامعة لندن. بالإضافة إلى نقل المقياس إلى إيطاليا وألمانيا وروسيا والسويد واليابان في عام 1920 ترجم أحمد فكري مقياس بينيه الأصلي إلى اللغة العربية. وقدم له تعليمات تفصيلية، وكان يعتقد خطأ أن للمدرس - دون تدريب سابق - يمكنه تطبيق المقياس والحكم على نكاء الطفل. ومع ذلك فإن أكثر هذه المحاولات شهرة في تاريخ لتقياس النفسي محاولة لويس ثرمان (1877-1956) أستاذ علم النفس بجامعة ستانفورد، والتي حرف بها المقياس باسم مقياس ستانفورد بينيه للنكاء.

وقد ظهرت الطبعة الأولى لهذا المقياس عام 1916. وأدخل ترمين على المقياس الأصلي عدداً من التعديلات والإضافات بحيث يمكن القول أنه اختبار جديد، فقد كان حوالي ثلث الأسئلة جديداً، وحل ترمين عدداً من الأسئلة القديمة وأعيد توزيعها على المستويات العصرية المختلفة أو حذفت. وأعيد تقنين المقياس كله على عينة من الأمريكيين بلغ عددها 1000 طفل و400 راشد. وأعدت تعليمات مفصلة لتطبيق كل اختبار وتصحيحه، واستخدام معيار لمبة الذكاء (تصديلاً لمعيار النسبة العقلية لشترن وكولمان).

وقد كتبت مجلة المقطف المصرية مقالاً عن هذا المقياس عام 1927 كما كتب عنه الدكتور أمير بقطر عام 1928، ونقله إلى اللغة العربية لأول مرة طبيب خير متخصص في علم النفس هو الدكتور حسن عمر.

وفي عام 1937 ظهرت الطبعة الثانية المصححة للمقياس، وتتكون من صورتين متكافئتين، الصورة (أ) والصورة (ب)، وفي هذه الطبعة زاد عدد أسئلة المقياس وأعيد تقنيته تماماً على عينة جديدة جيدة الانتقاء من المجتمع الأمريكي بلغ عددها أكثر من 3000 مفحوص تتراوح أعمارهم بين  $1\frac{1}{2}$  إلى سن 18. واشتملت كل عينة عصرية على عدد متساو من الذكور والإناث. وتم اختبار معظم المفحوصين من سن 5 سنوات وما بعدها من المدارس، أما عينات لطفال ما قبل المدرسة (عينات الأعمار  $1\frac{1}{2}$ ، 2، 2.5، 3، 3.5، 4، 4.5، 5 سنة) فلم الحصول عليها بطرق مختلفة، أيسرها تطبيق المقياس على أخوة لطفال وتلاميذ المدارس، وقد روعي في انتقاء العينات للتوزيع الجغرافي والمستويات الاقتصادية والاجتماعية.

وفي عام 1960 ظهرت طبعة ثالثة لمقياس ستانفورد بينيه في صورة واحدة فقط تجمع الفضل الأسئلة من صورتين طبعة عام 1937، وبدون تقويم أي محتوى جديد أو صياغة أسئلة جديد فكان من الممكن حذف الأسئلة التي لا تتناسب العصر، وأبعد توزيع الأسئلة التي تغيرت مستويات صعوبتها في الفترة بين التحليلين تبعاً للتغيرات الثقافية، وفي إعداد طبعة عام 1960 ولجّه ترمان مشكلة شائعة في القياس النفسي وتتلخص في أنه من المستحب عامة لإخال تعديلات دورية على المقاييس للإفادة من التقدم العلمي في ميدان بناء الاختبارات، ومن الخبرة السابقة باستخدام الاختبار، بالإضافة إلى جعل محتوى الاختبار مساهراً لروح العصر. وتردد الحاجة إلى الاعتبار الأخير بالنسبة لأسئلة للمعلومات والمواد المصورة والتي تتغير تبعاً لتغيير الظروف الثقافية، فاستخدم أسئلة قديمة يؤثر في علاقة الناحية بالمفهوم وفي الصدق الظاهري للاختبار قد ينير في مستوى صعوبة الأسئلة. ومن ناحية أخرى فإن التعديل قد يجعل معظم البيانات المتولدة عن الاختبار لا ترتبط بالصورة الجديدة، فالاختبارات التي استخدمت استدلالات واسعة لعدة سنوات كمقياس ستانفورد بينيه تتأثر عنها ثروة من المعلومات يجب أن توزن في مقابل الحاجة إلى التعديل. ولهذه الأسباب فضل تلاميد ترمان الجمع بين الصورتين في صورة واحدة بعد اللقاء لفضل ما فيهما. وإعداد صورة مكافئة ليس ثمة فائدة لتحقيق هذه الغاية، وخاصة أنه أصبحت توجد في الوقت الحاضر اختبارات فردية أخرى جيدة للصنع لم تكن متولدة عام 1937.

#### تعديل اختبارات المقياس

لقد استغرق تعديل عام 1937 لمقياس ستانفورد بينيه أكثر من عشر سنوات من البحث العلمي المنظم حول تجمع الأسئلة الجيدة وتجهيزها على جامعات صغيرة للحد من المفوضين الذين طبق عليهم مقياس عام 1916.

وكانت الخطوة الأساسية في ذلك هي تطبيق الصورتين المبدئيتين في ذلك الوقت على عينة الفتيان (3184 مفوضاً) ومنها تم الانتقاء النهائي للأسئلة وتوزيعها على الأعمار وقد أعطيت لكل مفوض صورتا الاختبار بفواصل زمنية يتراوح بين يوم واحد وأسبوع واحد. وفي هذه الخطوة كان معك تحليل الأسئلة لتحديد صحتها هو العمر الزمني والدرجة الكلية للمركبة من الصورتين المبدئيتين. ويرتبط المعك لثاني بما أسماه الاتساق الداخلي حتى يتحقق في المقياس درجة من التجانس. واستخدمت ثلاثة مقاييس في تحليل كل سؤال (الخيار) في المقياس هي:

- 1- منحنيات النسب المئوية للمفوضين الذين يجتازون السؤال بنجاح في الأعمار الزمنية المتتالية.
- 2- منحنيات للنسب المئوية للمفوضين الذين يجتازون السؤال بنجاح في الفئات المتتالية للدرجة الكلية في المقياس (وبالطبع لكل صورة من صورتين المقياس منحنياتها الخاصة).
- 3- معامل الارتباط اللثاني بين كل سؤال (الخيار) والدرجة الكلية في المقياس.

ويتفق الطريقة الأولى مع أسلوب بينيه الذي كان يستخدمه في تحليل المفردات ولانقاء الأسئلة التي تصلح للمستويات العمرية المختلفة. وفي هذا يجب أن نشير إلى أن ترميزاً استخدم للحكم على صلاحية الأسئلة نسباً مئوية في الأعمار الصغيرة أعلى منها في الأعمار الكبيرة. ومن ذلك مثلاً أنه كان يعتبر السؤال مناسباً لمن 3 سنوات إذا أجاب عليه 73% من أطفال هذا السن إجابة صحيحة. أما في حالة الأطفال من سن 10 سنوات فكانت هذه النسبة 59%.

ويشير ملاحظتنا إلى أن هذا الاختلاف في التنبؤ للموتية هو من المطالب للضرورة في مقاييس العمر. فمن الملاحظ أن الانحراف المعياري للعمر العقلي يجب أن يتزايد مع العمر إذا كان لا بد للنسبة للكفاءة أن تظل ثابتة. وعلى ذلك فإذا كان التباين في العمر العقلي في الأصغر عقليا أكبر منه في الأصغر للنسبة فإن ذلك يعني ضرورة وجود انتشار أكبر لأداء المفحوصين الأكبر في مستويات الأصغر المتجاورة. وعلى ذلك فمع الزيادة في العمر يقل عدد المفحوصين الذين يجتازون بنجاح. السؤال المطروح لمصنوعي المقاييس.

ويمكن أن نوضح هذه المسألة بمثال بسيط لنفرض أنه بين الأطفال من سن 5 سنوات يصل 40% منهم إلى العمر العقلي 5، بينما يصل 30% إلى أعلى، و 30% آخرون إلى أعمار أقل من هذا المستوى. وحيث أن أي مفحص عمره العقلي 5 سنوات أو أكبر يجب أن يجتاز من الوجهة النظرية السؤال الخاص بـ 5 سنوات فإن 70% (40% + 30%) من الأطفال من فئة 5 سنوات فتوقع نجاحهم في أسئلة سن 5 سنوات. ولنفر من ناحية أخرى أنه من بين الأطفال من سن 10 سنوات لا يصل إلى العمر العقلي 10 سنوات إلا 20%، بينما يصل إلى أعمار أعلى من ذلك 40% منهم، وإلى أعمار أقل من ذلك 40% آخرون. وبالطبع لا يمكن الحصول على مثل هذا التوزيع إلا إذا كان التباين في العمر العقلي للأطفال من 10 سنوات أكبر منه في 5 سنوات. ومعنى ذلك أنه من بين الأطفال من سن 10 سنوات تصبح النسبة المتوقعة لاجتياز الأسئلة المخصصة لهذا العمر هي (20% + 40% - 100%). وفي هذا المثال تصبح النسبة للموتية التي يمكن استخدامها معاً لانتقاء الأسئلة في سن 5 سنوات هي 70%، وفي سن 10 سنوات هي 60%.

وهي الممارسة الواقعية لانتقاء اختبارات مقياس ستانفورد بيهيه نجد أن المحك يتناقص من 77% عند مستوى من عامين إلى ما هو أقل قليلاً من 50% في مستوى الرشد العادي. ونال النسب عن هذا في انتقاء أسئلة مصفويات الرشد المتتوق وذلك للوصول إلى سقف مناسب للاختبار، وعند الانتقاء النهائي للمفردات لم يقتصر الاهتمام على تمايز العمر والاتساق الدلالي لمصعب، وإنما امتد ليشمل أيضاً مدى تناقص الفروق بين الجنسين وتوازنها في النسب المثوية للناجحين في مختلف المستويات العمرية . واستبعدت بالفعل المفردات التي اختلف فيها أداء أحد الجنسين عن الآخر بنسب لها دلالة إحصائية، على أساس أن مثل هذه المفردات قد تكل على تباين لفظي أو غير مرتبط بما يقيسه الاختبار نتيجة للخبرات السابقة للجنسين.

ولي إعداد طبعة عام 1960 من المقياس لاختبار الأسئلة من الصورتين له، م على أساس أداء 4498 مفحوصاً تتراوح أعمارهم بين عامين، 18 سنة الذين أجلبوا على صورتين المقياس أو إحداهما بين علمي 1950، 1954. كما تضمنت الحنة مجموعتين طبقيتين لشملاً على 100 طفل من سن 6 سنوات، و100 مراهق من سن 15 سنة، صنفنا حسب مهنة الوالد ومسئول التعليمي.

وبعني هذا أن طبعة 1960 لم تتضمن إعادة لتقنين المقياس. واستخدمت العينات الجديدة لمراجعة التغيرات في صعوبة الأسئلة فقط بعد ما تحددت مصفويات للصعوبة بإيجاد النسب المثوية للأطفال الذين يجتازون السؤال في أعمار عقلية متتالية في صورتين عام 1937. كما روجعت بعض الأسئلة للتحقق من احتمال وجود فروق جغرافية أو اقتصادية اجتماعية وذلك من خلال

المقارنة بين المجموعات الفرعية التي تكون العينة الكلية. ولم يتضمن المقاييس كما أُرنا، أي مواد جديدة، وإنما عدلت بعض الرسوم الخاصة بالأشياء الشائعة معاصرة للعصر. وفي هذه الأحوال اختبرت الاختبارات المطلة قليلاً على عينات صغيرة قبل تضمينها في المقاييس.

### وصف المقاييس

يتألف مقياس متلفرد بينيه من صندوق مجموعة من اللعب المغطاة تستخدم مع الأعمار الدنيا، وكثيرين من البطاقات المطبوعة، وكراسة لتسجيل الإجابات، وكراسة للتعليمات، وكراسة سجلير التصحيح.

وقد رُتبت أسئلة (اختبارات) للمقياس وجمعت تبعاً لمستويات الأعمار التي تمتد من سن عامين حتى سن الرشد المتفوق. وتتحدد مستويات لمقياس من سن 2 حتى سن 5 سنوات في ثلاث نصف سنوية على النحو الآتي:

2، 6-2، 3، 4-3، 6-4، 5 (مع ملاحظة أن الرمز 6-2، 6-3، 4-6، 4-6، يدل على من عامين و6 شهور، 3 سنوات و6 شهور، 4 سنوات و6 شهور على التوالي) وذلك بسبب التقدم السطي السريع في الأعمار الدنيا. أما الفترات من سن 5 سنوات حتى 14 سنة فتتحدد مستويات الأعمار فيها في ثلاث سنوية، وتكفل للمستويات الأخرى في المقاييس على الرائد المتوسط والرائد المتفوق. وكان عدد أسئلة (اختبارات) كل مستوى عمري 6 اختبارات، باستثناء مستوى الرائد المتوسط الذي يبلغ عدد أسئلته 8 اختبارات.

وكافلت الأسئلة في كل مستوى عمري من مستوى صعوبة موحدة تقريباً، وثرُتب بصرف للنظر عن أي فروق طفيفة في الصعوبة. وقد أعد سؤال



الاحتياطي في كل مستوى عمري معادل في صعوبته لأسئلة العمر. ويستخدم هذا السؤال الاحتياطي إذا تطلب الأمر استبعاد أحد الأسئلة العادية بسبب بعض الظروف الخاصة تجعله غير ملائم للمفحوص، أو بسبب عدم الالتزام في التطبيق بحرمة التعليمات بما يتعارض مع تقنين المقياس. وهكذا يبلغ العدد الكلي لاختبارات المقياس 142 اختباراً (122 اختباراً أصلياً، 20 اختباراً احتياطياً).

وقد اختير 4 أسئلة من كل مستوى عمري على أساس الصدق والتمثيل ليتألف منها المقياس المختصر الذي يستخدم حين لا يسمح الوقت بإعطاء المقياس الكلي، وتؤكد المقارنات بين نسب الذكاء كما تتحدد بالمقياس الكلي والمقياس المختصر تطابقاً كبيراً بينهما، بحيث يصل معامل الارتباط بينهما إلى مستوى معامل ثبات المقياس الكلي.

### تطبيق المقياس

يتطلب هذا المقياس، كغيره من اختبارات الذكاء الفردية، فاحصاً على درجة كبيرة من المهارة والتدريب لأن إعطاء اختباره وتصحيحها على درجة كبيرة من التقيد ويستلزم أن تتوفر في الفاحص آفة شديدة بها وخبرة طويلة في التطبيق، لأن تردد الفاحص أثناء التطبيق قد يؤدي إلى فقدان الملاقة المطلوبة بينه وبين المفحوص، بل وفقدان ثقة المفحوص به. بل إن أدنى تغيير في صياغة الأسئلة قد يؤدي إلى تغيير في مستوى صعوبتها كما تتحدد في عينه للتقنين. بالإضافة إلى ما يتطلبه المقياس من تصحيح فوري لاختباره عقب إعطائها مباشرة، لأن تتابع إعطاء الاختبارات يستدعي أداء المفحوص في المستويات السابقة.

ويرى كثير من علماء النفس الإكلينيكيين أن مقياس ستانفورد بينيه ليس مقياساً مقنناً للكفاءة فحسب، وإنما هو كذلك نوع من المقابلة الإكلينيكية، لأنه يهيئ فرصة للتفاعل بين الفاحص والمفحوص، ويهيئ مصادر أخرى للدلالات الإكلينيكية لدى الأخصائي الإكلينيكي المتمرس. بل إن مقياس ستانفورد بينيه، أكثر من غيره من الاختبارات الفردية، يهيئ الفرصة للفاحص أن يلاحظ الطرق التي يلجأ إليها للمفحوص في العمل، وأساليب حله للمشكلات، وغير ذلك من الأساليب الكيفية للأداء، بل إن الفاحص يمكنه أن يحكم من خلال الموقف الاختباري الذي يستخدم فيه هذا المقياس على بعض خصائص الشخصية مثل مستوى النشاط والثقة بالنفس والمثابرة والقدرة على التركيز. وبالطبع فإن أي ملاحظات كمية يمكن تسجيلها أثناء موقف الاختبار يجب تحديدها في ذاتها، ولا يجب تفسيرها على نفس النحو الذي تفسر به درجات الاختبار الموضوعي. وتعتمد قيمة هذه الملاحظات الكيفية إلى حد كبير على مهارة الفاحص وخبرته وإعداده السيكلوجي الصيق، بالإضافة إلى وعيه الشديد بمزلق وحدود هذا النوع من الملاحظات المباشرة. وقد استطاع مورياتزي أن يسجل أنماط الملاحظات الإكلينيكية في موقف الاختبار الفردي، على أساس أن جملة الاختبار هي فرصة لدراسة سلوك الطفل أثناء مواجهة موقف تسود فيه عوامل التحدي والامتناع والصعوبة والإحباط.

ولم يإطاء مقياس ستانفورد بينيه لا يحاول المفحوص جميع أسئلة المقياس، وإنما يختار للمفحوص فقط حول مدى المستويات العمرية الثلاثة لمستواه العقلي. ولا يتطلب المقياس أكثر من 30-40 دقيقة للأطفال الصغار، ولا يتطلبه أكثر من ساعة ونصف للراشدين. والإجراء المقنن هو أن تبدأ الاختبار من مستوى أدنى قليلاً من العمر العقلي المتوقع للمفحوص وعلى ذلك

فإن الاختبارات الأولى يجب أن تكون سهلة بالقدر الذي يستلزم في المفحوص  
 ثقة بالنفس، بشرط ألا تكون سهلة جداً بحيث تؤدي بالمفحوص إلى الاستهانة  
 بالاختبار أو للامبالاة. فإذا فشل المفحوص في أي اختبار من اختبارات المستوى  
 الذي يبدأ به يعطى اختبارات المستوى الأدنى منه، وهكذا يستمر الفحص حتى  
 يصل إلى المستوى الذي يستطيع المفحوص أن يجيب على جميع أسئلته،  
 ويصبح هذا المستوى هو ' العمر القاعدي '، ثم يعطيه الاختبارات العليا حتى  
 يصل إلى المستوى الذي يفشل المفحوص في الإجابة على جميع أسئلته ويسمى  
 في هذه الحالة ' الحد الأعلى للعمر '، وحين يصل الفحص إلى هذا المستوى  
 يتوقف عن إعطاء للمقاييس.

### تصحيح المقاييس

تصحح جميع أسئلة (اختبارات) مقياس ستانفورد بيليه على أسس النجاح  
 أو الفشل (واحد أو صفر). وتحدد كرامة التعليمات الأداء المطلوب ' للنجاح '  
 في كل سؤال. وقد تظهر نفس الأسئلة في مستويات صعبة مختلفة ولكنها  
 تصحح بمستويات مختلفة للنجاح وفي هذه الحالة تطبق الاختبارات مرة واحدة  
 ثم يحدد أداء المفحوص المستوى العمري الذي يستحقه. ومن ذلك مثلاً أن  
 اختبار المفردات يطبق ابتداء من المستوى العمري 6 سنوات حتى مستوى راشد  
 المتفوق (الثالث) اعتماداً على عدد للكلمات التي يستطيع المفحوص تحديدها  
 وتعرفها.

والواقع أن الأسئلة التي ينجح فيها المفحوص أو يفشل تظهر مقداراً من  
 التشتت في المستويات العمرية المتتالية. فمن لا نجد مفحوصاً يجيب على  
 جميع أسئلة عمره العقلي أو أدنى منه ويفشل في جميع الاختبارات الأعلى من

هذا المستوى؛ ولما نجد في العادة أن الاختبارات التي ينجح فيها تنتشر في مستويات عمرية عديدة يحددها العمر القاعدي من ناحية والحد الأعلى للعمر من ناحية أخرى. ويحسب العمر العقلي في هذه الحالة بالعمر القاعدي مضاعفاً إليه أي شهور عقلية لأي اختبار اجتازه من مستويات أعلى من عمره القاعدي.

### تفسير المعايير

من أهم مزايا مقياس ستانفورد بينيه ما يتوافر حوله من بيانات تفسيرية وخبرة إكلينيكية تراكمت طوال نصف قرن. وقد أصبح هذا المقياس عند كثير من علماء النفس الإكلينكيين والتربويين وغيرهم من المهتمين بالتقويم النفسي مرافقاً للنكاه. وأصبح تحديد مستويات النكاه يعتمد في جوهره على نسب النكاه كما يحددها المقياس، رغم ما في هذا الرأي من غلو لا يتسع المقام لمناقشته.

والواقع أن شيوع تصنيف مستويات نسب النكاه، رغم فائته في كثير من تفسير الأداء الاختباري، يتضمن بعض المخاطر. فهو كغيره من صور التصنيف الملوكي يجب أن يستخدم بمرونة كافية، كما يجب ألا يستبعد البيانات الأخرى التي تتوافر عن المفحوص. وبالطبع لا يوجد حد فاصل تماماً بين فئات التصنيف. ومن ذلك مثلاً لا يوجد خط يميز تماماً بين التخلف العقلي والحالات الهامشية، أو بين التفوق والتفوق الشديد. فقد لوحظ مثلاً أن بعض المفحوصين من ذوي نسب النكاه 60 يتكيفون تكيفاً عادياً لمطالب الحياة اليومية، بينما نجد بعض الأشخاص من ذوي نسب النكاه 100 يتطلبون الإيداع في المؤسسات. كما، بعض الأشخاص من ذوي نسب النكاه 160 قد لا يحضرون حياة مثيرة، بينما نجد البعض الآخر من ذوي نسب النكاه الأقرب إلى 100 يسهمون بعض الإسهامات الممتازة. ومعنى ذلك أن القرارات التي تتحدد حول تشخيص

النسب العقلي أو التفوق العقلي يجب ألا تعتمد على نسب الذكاء وحدها، وإنما يجب الاهتمام في حالة النصف العقلي بتغيرات أخرى مثل النصح الاجتماعي والتوافق الانفعالي والتظروف الصحية والجسدية والأحوال الأسرية والظروف المدرسية وغير ذلك من العوامل التي تؤثر في الأداء العقلي. وبالمثل فإن نسب الذكاء الحالية لا تترادف للمعقولة، لأن الإنجاز الرفع قد يتطلب بالإحسنة إلى ذلك الابتكار والمواهب الخاصة والمثابرة والاهتمام بالهدف وغيرها من العوامل العقلية والانفعالية والدافعية.

### ثبات المقياس

أجريت دراسات لثبات على طبعة 1937 من المقياس باستخدام طريقة الصور المتكاثفة على فترات زمنية مقدارها أسبوع أو أقل (للتحقق من مصدر ثبات الخطأ). الاستقرار الوقي وتكاثر المحتوى. وقد لوحظ أن المقياس يميل إلى أن يكون أكثر ثباتاً في الأعمار الكبيرة منه في الأعمار الصغيرة، وفي مستويات نسب الذكاء الدنيا عنه في نسب الذكاء العليا.

والواقع أن زيادة الثبات مع الزيادة في العمر هي من خصائص الاختبارات بوجه عام، وتتبع جزئياً من التحكم الأفضل في ظروف إعطاء الاختبار في حالة المفحوصين للكبار الص (وبخاصة حين تقارنهم بالأطفال ما قبل المدرسة). والعوامل الأخرى المساهمة في زيادة معاملات الثبات بظم معدل النمو مع التقدم في العمر.

لما عن ارتفاع معاملات الثبات في حالة نسب الذكاء الدنيا في أي عمر زمني اقترنبط فيما يبدو بخصائص مقياس متكافرد ببلية على وجه الخصوص.

فقد أشرنا إلى أن الأوزان النسبية لأسئلة الأعمار المختلفة تختلف، فهي في الأعمار الدنيا مقدارها شهر، وفي الأعمار المتوسطة شهران، وفي المستويات العليا قد تكون 4 شهور أو 5 شهور أو 6 شهور. ومعنى ذلك أن طريقة وزن الدرجات ميل إلى تضخيم أخطاء القياس في المستويات العليا، لأن النجاح أو الفشل للناجم عن الصنفة في سؤال واحد يؤدي إلى فرق أكبر في الدرجة للكلية في هذه المستويات عنه في المستويات الدنيا. وحيث أن المفحوصين من نوعي نسب الذكاء العالية (من أي عمر زمني) عادة ما يختبرون في المستويات العليا من المقياس، فإن نسب ذكائهم نتيجة لذلك تتضمن مقداراً أكبر من أخطاء المقياس وبالتالي تتخفض معاملات الثبات.

وقد حسبت معاملات الثبات لفئات مختلفة من مدى نسبة الذكاء في طبعة 1960، لأنه لوحظ أن جدول التنتار درجات الصورتين (م، ل) من طبعة 1937 يأخذ شكل المروحة، أي أن مدى التشتت يتضائل كثيراً في المستويات الدنيا، أما المستويات العليا فتظهر تشتتات واختلافات واسعة.

وبوجه عام يمكن القول أن مقياس ستانفورد بينيه على درجة عالية من الثبات، بالرغم من تجانس مجمرعات عينة التقنين من حيث العمر الزمني، مما يؤثر بالطبع في معاملات الارتباط عموماً ويؤدي إلى نقصانها. وقد تزلزلت هذه المعاملات بالنسبة للأعمار  $2\frac{1}{2}$  - 5 83 (ونسب الذكاء 140-149) 0.91 (النسب الذكاء 60-69). وبالنسبة للأعمار من 6-3 تزلزل بين 0.91، 97، وللفص للمستويات من نسب الذكاء، وبالنسبة للأعمار من 14-18 كانت معاملات الثبات المطلقة لمستويين هي 0.95، 0.98.

## صدق المقاييس

تتوافر بيانات كافية عن الصدق المرتبط بالمسكات سواء كان صدقاً تلامزياً أو صدقاً تنبئياً في ضوء معك التحصيل المدرسي. لمذ ظهر مقاييس ستانفورد بينيه عام 1916 لم يتوقف الباحثون عن حساب معاملات الارتباطات بينه وبين المدرجات المدرسية وتقديرات المعلمين ودرجات الاختبارات التحصيلية المقتلزة ويتراوح معظم هذه المعاملات بين 0.40، 0.75 كما ترتبط بالتقدم والتخلف المدرسي بوجه عام.

ولمذ وجد أن هذا المقاييس، كغيره من مقاييس الذكاء، يرتبط ارتباطاً حلياً بالأداء في جميع المواد الدراسية الأكاديمية تقريباً، إلا أن ارتباطه أعلى ما يكون بالمقررات اللغوية (كمواد للذات أو المواد الاجتماعية). ولمذ تأكدت هذه النتائج عند اختبار التلاميذ والطلاب في مختلف المراحل التعليمية، وبالطبع فإن معاملات الارتباط تميل إلى التناقص مع ارتفاع السلم التعليمي بسبب عوامل الانتقاء وعدم ملائمة صنف الاختبار مما يؤدي إلى نقصان التباين بين المفحوصين.

وفي تفسير نسبة الذكاء يجب أن نؤكد أن مقاييس ستانفورد بينيه، كغيره من اختبارات الذكاء العلم، هو في الأغلب مقاييس للاستعداد المدرسي أو الأكاديمي العلم، ومشبع إلى حد كبير بالمحتوى اللفظي، وبخاصة في اختبارات المستويات الصرية العليا. ولذلك فإن الأفراد من ذوي النفاذ اللغوية، أو أولئك الذين تتوافر لديهم قدرات غير لغوية يحصلون على درجات منخفضة نسبياً في المقاييس. وكذلك فإنه توجد بعض المجالات لا يلعب فيها الاستعداد المدرسي أو القهم اللغوي دوراً هاماً، وفيها لا تصلح اختبارات الذكاء من نمط ستانفورد بينيه.

ومعنى ذلك أن استخدام الاختلافات - بصفة عامة - يجب أن يكون في المواقف التي تلائمها، وإلا فإن الباحث يتوقع من المقياس ما لا يطمح.

### صدق التكوين الفرضي

من أهم المحركات التي استخدمت في انتقاء لاختبارات مقياس ستانفورد بينيه محك تمايز العمر. ومعنى ذلك أن المقياس يقيس قدرات متزايدة مع العمر منذ الطفولة حتى الرشد (في ظل إطار ثقافي وحضاري معين بالطبع).

وقد استخدم أيضاً محك الاتساق الداخلي في انتقاء الأسئلة. وقد كشفت النتائج وجود نوع من التجانس الوظيفي في مقياس ستانفورد بينيه، بالرغم من التنوع الظاهري للمحتوى، والدليل على هذا أن متوسط معاملات الارتباط بين السؤال والمقياس هو 0.66 في طبعة عام 1960 إلا أن من الواضح غلبة الطابع اللغوي على المقياس، ويتضح هذا من ارتفاع معامل الارتباط بين الأسئلة اللغوية والأداء الكلي في المقياس عنه في حالة الأسئلة غير اللغوية.

ومن البراهين الإضافية على صدق التكوين الفرضي نتائج الدراسات العلمية الحديثة التي أجريت على اختبارات مقياس ستانفورد بينيه. ويعتمد هذا البرهان على أنه لو كانت نسب الذكاء قابلة للمقارنة في مختلف الأعصار، فإن للمقياس يجب أن يكون له نفس التكوين العالمي في جميع مستويات الأعصار، بل يجب أن يكون للمقياس مثبناً بعامل واحد حتى يمكن تفسير نسبة الذكاء تفسيراً واضحاً. أما إذا كان المقياس مثبناً بعاملين مختلفين أو أكثر فإن نسبة الذكاء للولادة التي يحصل عليها أفراد متحدون (نسبة ذكاء 120 مثلاً) قد تقل على فترة معينة في بعض الحالات، وفترة أخرى في البعض الآخر.



وقد قام ماكتمر بإجراء عدة تحليلات عاملية منفصلة لاختبارات مقياس ستانفورد بينيه في 14 مستوى عمري (2، 2½، 3، 3½، 4، 4½، 5، 6، 7، 8، 9، 11، 13، 15، 18). وكل عدد الأفراد المستخدم في كل تحليل عاملي يتراوح بين 99، 200، كما تراوح عدد الاختبارات بين 19، 35. وفي كل الأحوال حُصيت معاملات الارتباط الرباعية بين المقدرات، وخضعت مصفوفات الارتباط للتحليل العملي. وتؤكد نتائج ماكتمر أن الأداء في مقياس ستانفورد بينيه يمكن تفسيره في ضوء عامل عام واحد، كما وجدت عوامل طائفية إضافية في قليل من المستويات العمرية. إلا أن إسهام هذه العوامل في التباين كان قليلاً. كما ثبت أن العامل للعام المشترك الذي وجد في المستويات العمرية المقاربة كان متشابهاً إلى حد كبير. ومن أهم النتائج التي توصل إليها ماكتمر أن العامل يصبح أكثر لفظية كلما تقدمت الأعمار. ومن ذلك مثلاً أن تشبع العامل للعام باختبار المقدرات ارتفع من 0.59 في سن 6 سنوات إلى 0.91 في سن 18 سنة.

لجري جونز دراستين علميتين منفصلتين لاختبارات مقياس ستانفورد بينيه على 4 مجموعات من الأطفال أعمارهم (7، 9، 11، 13) على التوالي، وكانت كل مجموعة تتكون من 100 ذكر، 100 أنثى. ووجد أنه في كل عصر يوجد عدد من القدرات المتشابهة ولكنها مرتبطة، ومنها العوامل اللفظية والانتكارية الاستدلالية والمكانية والإدراكية. وقد تختلف هذه العوامل أوزانها النسبية من مستوى عمري لآخر.

ولموقع أن الأساليب الإحصائية التي استخدمها ماكتمر وجونز من النوع الذي يؤكد دور العامل العام في الحالة الأولى (دراسات ماكتمر)، ودور العوامل

الطائفة المختلفة في الدراسات التي قام بها جولد. وحين نتأمل النتائج التي ترسل إليها قباحان نجد أنها تؤكد وجود قدر مشترك في القياس ككل، وهي خاصية من خصائص بناء القياس حسب طريقة انتقال أسئلته التي كانت تؤكد ارتفاع معاملات الارتباط بين السؤال والدرجة الكلية في القياس. وبالإضافة إلى ذلك فإن الأداء يتأثر كذلك بعدد من القدرات الخاصة التي يختلف تكوينها تبعاً لمستوى العمر الذي تختبره.

أما عن الدراسات التي أجريت في مصر لتحديد المعالم السيكمترية لمقياس ستانفورد بينيه فإنها قليلة. وتتلخص في دراسة قام بها الدكتور لويس كامل مملكة عام 1960 للتحقيق من صدق الترتيب المبدئي لاختبارات المفردات في القياس، وذلك بتطبيقها على مجموعات كبيرة من الأفراد ومن مختلف فئات السن (بين 3 سنوات، 30 سنة) بلغ عددها 1642 مفحوصاً وقد تكون للدراسة الوحيدة التي تناولت مقياس ستانفورد بينيه ككل بحث محمود الهولوي عام 1963 عن ثبات القياس وصدقته. وقد استخدم في حساب الثبات طريقة إعادة الاختبار على 95 تلميذاً بالسنه الثالثة الإعدادية (متوسط أعمارهم 14 عاماً) بفواصل زمني بين شهر وشهرين وبلغ معامل الارتباط 0.83 أما عينة للصنق فتألفت من 160 تلميذاً بالسنه الثالثة الإعدادية أيضاً، وقد حسب لهم معاملات الارتباط بين درجات القياس (نسبة الذكاء) وحالات الرسوب والنجاح في امتحان الشهادة الإعدادية من ناحية، والمجموع الكلي لدرجة الامتحان من ناحية ثانية، ودرجات مواد هذا الامتحان من ناحية ثالثة. وقد وجد هذا الباحث أن معامل الارتباط الفعلي بين نسب الذكاء وحالات النجاح والرسوب هو 0.54، ومعامل الارتباط بين نسب الذكاء والمجموع الكلي لامتحان الشهادة الإعدادية 0.818 أما معاملات ارتباط القياس بدرجات المواد الدراسية فكانت على النحو الآتي:

الصحة والطوم	0.596	الجغرافيا	0.509	اللغة العربية	0.534
التربية الوطنية	0.511	اللغة الإنجليزية	0.485	الفرنسية	0.489
التاريخ	0.487	الحبر	0.511	الحساب	0.557

وأعد الباحث بعد ذلك مصفوفة معاملات ارتباط درجات للنكاه ودرجات المواد الدراسية المختلفة وأخضعها للتحليل الفعلي فوجد أن تتبع مقول ستانفرد بييه بالعامل العام 0.784 وهو أعلى التشبهات بهذا العامل وفي هذا برهان مبدئي على صدق التكوين الفرضي للمقياس، وجد أن باقي المتغيرات كانت درجات التحصيل المدرسي، ولم تتضمن المصفوفة أي مقياس آخر للنكاه.

#### مقاييس وكسار للنكاه

ظهرت الطبعة الأولى من مقاييس وكسار والتي عرفت باسم مقياس وكسار بلغوي للنكاه عام 1929. وكان أحد الأغراض الهامة من إعداد هذه المقاييس توفير اختبار النكاه يصلح للراشدين. وقد أشار وكسار في ذلك الحين إلى أن الاختبارات التي كانت متاحة حتى ذلك الوقت، ومنها بالطبع مقول ستانفرد بييه، محدودة في جوهرها للأطفال وتلاميذ المدارس، وكانت تستخدم مع المراهقين والراشدين عن طريق إضافة بعض الأسئلة الأكثر صعوبة من نفس النوع. ولذلك فإن من المشكلات الهامة في هذه الاختبارات أن محتواها لا يثير اهتمام الراشدين وإذا لم يتوافر في المقاييس حداً أدنى من المسبق الظاهري يسمح تكوين علاقة طيبة بين الفاحص والمفحوص الراشد.

وبالإضافة إلى ذلك فإن المبالغة في تأكيد عامل السرعة في معظم الاختبارات الفردية يؤدي إلى تعطيل أداء الراشدين. وكذلك فإن الاهتمام بالتناول

الروتيني للتلويحي اللفظية تحظى بالاهتمام الأكبر في هذه المقاييس. ثم إنه يستحيل استخدام معيار العمر العقلي مع الراشدين كما نشرنا. وأخيراً فإن عينات تقنين الاختبارات الفردية مثل مقياس ستانفورد لم تتضمن إلا قليلاً من الراشدين.

ومواجهة هذه المشكلات أعد وكسلر مقاييسه للذكاء، ومن أشهرها مقاييسه للذكاء الراشدين. والذي يسمى في الوقت الحاضر Wechsler Adult Intelligence Scale WAIS. والذي ظهر في عام 1955 محل مقياس وكسلر السابق للتقديم. وبالطبع يتشابه المقياس الجديد في الشكل والمحتوى مع المقياس السابق. إلا أنه تتطلب على بعض الصعوبات اللفظية الخاصة بمدى تمثيل عينة التقنين. رتبنا الاختبارات الفرعية التي يتكون منها المقياس.

#### وصف مقياس وكسلر لذكاء الراشدين

يتكون مقياس WAIS من 11 اختباراً فرعياً، ستة منها تصنف في مقياس لفظي، 5 منها تصنف في مقياس أدائي أو عملي performance. و فيما يلي وصف موجز لهذه الاختبارات الفرعية مع بيان عدد الأسئلة التي يتكون منها كل اختبار في الطبعتين العربية والأجنبية.

#### المقياس اللفظي:

يتكون المقياس اللفظي من 6 اختبارات هي:

- 1- المعلومات العامة: information، و يتكون من 29 سؤالاً في الطبعة الأجنبية (26 سؤالاً في الطبعة العربية منها سؤال تمهيدي) تشمل مدى

واسعا من المعلومات العامة التي يفترض شيوعها في ثقافة الراشدين.  
وقد روعي تحرر الاختبار من المعارف التخصصية و الأكاديمية.

2- لفهم العام: Comphrnsion ، و يتكون من 14 سؤالا في الطبعة الأجنبية ( 10 أسئلة في الطبعة العربية )، و في كل سؤال يشرح المفهوم ما يجب عمله في ظروف معينة. ويتطلب هذا الاختبار قدرة على الحكم العملي و الفهم العام common sense، وهو يشبه في هذا أسئلة الفهم في مقياس متفرد بينيه، إلا أن محتواه أقرب إلى اهتمامات الراشدين و نشاطهم.

3- الاستدلال الحسابي: arithmetic، و يتكون من 14 مسألة حسابية في الطبعة الأجنبية ( 10 مسائل في الطبعة العربية ) تتفق مع مستوى مادة الحساب في المرحلة الابتدائية. و تعرض كل مسألة شفويا، و يحلها المفحوص شفويا كذلك دون استخدام الورقة و القلم، و لكل مسألة زمن محدد.

4- المتشابهات: Similarities، و يتكون من 13 سؤالا في الطبعة الأجنبية ( 13 سؤالا في الطبعة العربية ) تطلب من المفحوص أن يحدد كيف يتشابه شيان ( البرقاقة و الموزة مثلا ) .

5- مدى الأرقام ( أو إعادة الأرقام ) : digit span، و فيه تعرض على المفحوص قوائم تتكون من 3 أرقام إلى 9 أرقام، و يكون عليه إعانتها شفويا، و في القسم الثاني من الاختبار يكون على المفحوص إعادة الأرقام بالعكس.

6- للمفردات: vocabulary، و فيه تعرض على المفحوص شفويا و بصريا في وقت واحد 42 كلمة مترابطة الصعوبة، و يطلب منه تحديد معنى كل كلمة.

## المقياس الأولي أو الصل:

يتكون المقياس العملي من 5 اختبارات هي:

1-رموز الأرقام: digit symbol، و هو عبارة عن صورة من اختبار التعويض الضمري code substitution المألوف في مقاييس الذكاء غير اللغوية. و يتكون مقياس لشدة من 9 رموز تزوج مع الأرقام للتمسية. و يحلى هذا المفتاح للمفحوص و يكون عليه أن يكتب أكبر عدد من الرموز التي تثل على الأعداد الموجودة في ورقة الإجابة في زمن لا يتجاوز 1.5 دقيقة.

2- تكميل الصور: Picture completion، و يتكون من 21 بطاقة في الطبعة الأجنبية ( 15 بطاقة في الطبعة العربية ) كل منها يحتوي على صورة ناقصة، و يكون على المفحوص أن ينكر الجزء الناقص من كل صورة.

3- رسوم المكعبات: block design، و تشبه مواد هذا الاختبار ما يوجد في اختبار Kohs Block Design. و يتكون من صندوق به 16 مكعب، و 9 بطاقات على كل منها رسم مختلف ( اثنان من هذه البطاقات للتدريب ). وقد لونت ألوجه المكعبات بالألوان الآتية في الأصل الأجنبي: الأحمر فقط و الأبيض فقط و الأبيض\_الأحمر معا، أما في الطبعة العربية فهي ملونة بالأزرق فقط و الأبيض فقط و الأصفر فقط و الأبيض\_الأحمر معا والأزرق\_الأصفر معا. و يطلب من المفحوص إنتاج رسوم على الور الذي تحده البطاقات ذات مستويات

متريدة التمتع تطلب ما بين 4 مكعبات، 9 مكعبات، و لكل بطاقة زمن محدد.

4- ترتيب الصور: Picture arrangement، و يتكون كل سؤال فيه من مجموعة من البطاقات، عليها صور، و يطلب من المفحوص ترتيبها بالتتابع الصحيح بحيث تعطى قصة مفهومة. و لكل سؤال وقت محدد (عدد الأسئلة في الطبعة الأجنبية 8 وفي الطبعة العربية 6).

5- تجميع الأشياء: Object assembly، و يتبع هذا الاختبار نفس النمط المستخدم في اختبار Pinter\_Paterson الأدائي. و يتكون في الطبعة العربية من نمالاج من الخشب ثلاثة أشياء هي الصبي أو المانيكان، الوجه أو البروفيل، اليد، قطع كل منها إلى قطع مختلفة، و يطلب من المفحوص في كل منها جمع للقطع بحيث تكون الشكل الكامل.

### تصحيح المقاييس:

تصحيح اختبارات الاستدلال الحسابي و رسوم المكعبات و ترتيب الصور و تجميع الأشياء تبعاً لسرعة الأداء و نقتطع على النحو الآتي:

1- الاستدلال الحسابي: تعطى درجة واحدة لكل مسألة تحل حلاً صحيحاً في الوقت المحدد لها، وتعطى درجة واحدة إضافية للزمن في المسألتين 9، 10 إذا حلت كل منهما في حدود 40 ثانية (الزمن العادي لكل منهما 120 ثانية)، وتعطى درجتان إضافيتان إذا حلت المسألة في حدود 15 ثانية، يحسب الزمن بعد انتهاء المفحوص من قراءة المسألة (النهاية المظلمة 14).

2- رموز المكعبات: تقدر الرسوم حسب الدقة والزمن الأقصى، وتعطى درجات إضافية في حالة إتمام الرسم في زمن أقل. وقد أحد جدول لإعطاء تقديرات حسب زمن الإنجاز (النهاية العظمى 42).

3- ترتيب الصور: تقدر الدرجة في هذا الاختبار على أساس دقة الترتيب والزمن المستغرق فيه، وأعدت جدول لهذا الغرض بحيث تصحح كل من المجموعات الثلاث الأولى على أساس التقدير (صواب أو خطأ)، أما المجموعات الثلاث الأخيرة فتعطى تقديرات جزئية على كل ترتيب لا يطابق الترتيب الصحيح، ولكن يمكن قبوله. وبالنسبة للمجموعتين الأخيرتين تعطى تقديرات إضافية على الترتيب الصحيح إذا تم في زمن معين. وتحسب الدرجة النهائية في هذا الاختبار بجمع التقديرات الجزئية في المجموعات المختلفة (والنهاية العظمى 21).

4- تجميع الأخطاء: يصحح الشكل الأول (الصحي) حسب الدقة فقط، أما الشكلان الآخران (الوجه واليد) فيصحح كل منهما حسب الدقة والزمن معاً. وفي الحالة الأخيرة تكون الدرجة في مواد الاختبار هي مجموع الكلي للتقديرات على الدقة والزمن جميعاً. وتوجد بكراسة التعليمات بروتوكولات خاصة لتصحيح كل شكل حسب الدقة، وجدول لتصحيح على أساس الزمن (النهاية العظمى 26).

أما اختبارات المطومك العامة، ومدى الأرقام (أو إعادة الأرقام)، وتكسيل الصور، ورموز الأرقام فإن الدرجة الكلية فيها هي مجموع الإجابات الصحيحة حسب مفتاح التصحيح المبين كراسة التعليمات، والنهايات العظمى لهذه الاختبارات الأربع على التوالي هي: 25، 17، 15، 67.



لما اختبارات الفهم العام، والمتشابهات، والمفردات، فقد قام الدكتور لريس كامل ملوكة بإعداد نماذج التصحيح الخاصة بهذه الاختبارات الثلاثة ونشرها في كراسة منفصلة عام 1960، وفي هذه النماذج تعطى الأسئلة لاختباري الفهم العام والمتشابهات درجتين أو درجة واحدة أو صفر حسب درجة التصحيح في الإجابة ولوعه، وهكذا فإن النهاية العظمى لكل من هذين الاختبارين هي: 20، 24 على التوالي. أما اختبار المفردات فإن أسئلته تعطى في نماذج التصحيح درجة واحدة أو نصف درجة أو صفر (النهاية العظمى 42).

### المعايير وإجراءات التصحيح

لقد كانت هيئة التقنين المستخدمة في مقولس وكسلر للراشدين من المعينات للقلائل التي يتوفر فيها شرط التمثيل، فقد كانت تتكافئ من 1700 شخصاً من الرجال والنساء موزعين في مستويات عمرية صاعدة مبع بين 16 سنة و 64 سنة وأضيف إلى كل مستوى عمري رجل واحد وامرأة واحدة من المودعين بمؤهلات ضعيف العقول، وكذلك هيئة من الممثلين عددها 475 شخصاً تريد أعمارهم على 60 سنة.

وقد حولت الدرجات الخام في الاختبارات الفرعية إلى درجات معيارية متوسطها 10 وانحرافها المعياري 3. وكانت هذه الدرجات للموزونة مشتقة من جملة مرجعية عددها 500 حالة شملت جميع المفحوصين من سن 20 سنة، 34 سنة في هيئة للتقنين. وهكذا أصبحت الدرجات الفرعية في صورة وحدات قليلة للمقارنة، ثم حسبت درجات المقولس اللفظي، والمقولس الأدائي، والمقولس الكلي بجمع الدرجات للموزونة في الاختبارات اللفظية الست، والاختبارات الأدائية الخمس، والاختبارات الأحد عشر جُمعاً على التوالي. ثم حولت هذه

الدرجات الثلاثة إلى نسبة ذكاء الحرفية متوسطها 100 وانحرافها المعياري 15 في كل مجموعة عمرية، وهكذا يتحدد الوضع النسبي للمفحوص بمقارنة بالأشخاص من نفس مستواه العمري.

وقد استخدمت نفس الإجراءات في تقنين الطبعة العربية حيث استخدمت في إعداد الدرجات الموزونة ثلاث فئات عمرية هي:

- (1) من 20 إلى أقل من 25 سنة وعدد أفرادها 114.
- (2) من 25 إلى أقل من 30 سنة وعدد أفرادها 112.
- (3) من 30 إلى أقل من 35 سنة وعدد أفرادها 102.

وقد استخدمت هذه العينة البالغ عددها 328 مفحوصاً في إعداد نماذج ومعايير التصحيح لاختبارات الفهم والمشتبهات والمفردات بالإضافة إلى إعداد جدول للدرجات الموزونة وجدول لنسب الذكاء الحرفية لكل فئة من فئات العمر.

ولتصحيح هذا القياس نحلى بالمفحوص درجات ثلاث. نسبة الذكاء للكلية ونسبة الذكاء اللفظية ونسبة الذكاء العملية. ويحلى في الطبعة الأجنبية درجة رابعة تدل على ما يسمى وكسلر " معامل للكفاءة " أو (E.Q) ويتحدد بتقدير نسبة ذكاء الفرد في الفئة من 20 إلى 24 (في اختبار وكسلر بغير طبعة 1939) وفي الفئة من 29-75، بصرف النظر عن العمر الحقيقي بالمفحوص. ويستخدم هذا المعامل على أساس أن الفئة من 20-24 والفئة من 25-29 هما اللتان وصلتا إلى أعلى مستويات الأداء في القياس. ومعنى ذلك أن معامل الكفاءة طريقة الحكم على أداء الفرد في ضوء متوسط أداء الأفراد من المستوى العمري الذي يصل إليه الأداء الاختباري قصاء.

ويذكر التذكور لويس كليل ملبكة أنه يمكن الحصول على " معاملا كفاءة " بصورة مؤقتة عن طريق استخدام جداول نسب الذكاء للغة من 20 2، من العينة المصرية آفة الذكر، لأن هذه اللغة حققت بالفعل أعلى مستوى من الأداء في القياس.

### المعيار

أحد وكسلر عام 1944 فالت الذكاء التي يمكن أن تستخدم في أوضاع تصنيف الإكلينيكي، وكان أساس التصنيف هو حدود الخطأ المعياري المتوسط للشخص الذي تكون درجته أقل من المتوسط بمقدار يبلغ ثلاثة أمثاله للخطأ المعياري أو أكثر بعد من فئة ضعافت العقول، أما الذي يحصل على درجة تتراوح بين خططين معياريين وثلاثة أخطاء معيارية فيعد من قبيل الحالات الهامشية هكذا. ويبين الجدول الآتي أسس تصنيف مستويات الذكاء في مقياس وكسلر.

### جدول تصنيف مستويات الذكاء في مقياس وكسلر

التصنيف	جدول الخطأ المعياري	حدود نسبة الذكاء الانحراف	النسبة المئوية في السكان
ضعف العقول	3- فأقل	65 فأقل	2.2
الحالات الهامشية	من 2- حتى 3-	من 66 إلى 79	6.7
الأغبياء	من 1- حتى 2-	من 80 حتى 90	16.1
المتوسطون	من 1- حتى 1+	من 91 حتى 110	50.0
أعلى من المتوسط	من 1+ حتى 2+	من 111 حتى 119	16.1
الممتازون	من 2+ حتى 3+	من 120 حتى 127	6.7
الممتازون جداً	من 3+ فأكثر	من 128 فأكثر	2.2

## ثبات المقياس

حسبت معاملات الثبات للطبعة الأجنبية من مقياس وكسلر للنكاه الراشدين لفئات الأعمار 18-19، 24-34، 45-54 اختبرت من عينة التفتين على أساس تمثيلها لدى العمر الذي تشملته هذه العينة. واستخدمت طريقة التجزئة النصفية (بعد تصحيح سبيرمان- براون) في حساب ثبات كل اختبار فرعي فيه عدا اختبائي رموز الأرقام ومدى (إعادة) الأرقام حيث استخدمت معهما طريقة الصور المتكافئة لعدم صلاحية طريقة التجزئة النصفية لهما بسبب غلبة عامل السرعة للتدوين عليهما.

وبلاحظ أن معاملات الثبات للمقياس الكلي قد بلغت في المجموعات الثلاثة 0.97، وبلغت في المقياس اللفظي 0.96، وفي المقياس الأدائي 0.94 ومعنى هذا أن المقياس الثلاثة على درجة عالية من الثبات في ضوء الاتساق الدلطي.

أما في الطبعة العربية فقد حسب الدكتور فرج عبد القادر طه ثبات المقياس بطريقة إعادة الاختبار وطريقة التجزئة النصفية، وفي الطريقة الأولى ليعد تطبيق للمقياس على 40 حالة من العمال فتراوحت معاملات الثبات بين 0.6، 0.9 بالنسبة لكل من نسب النكاه والاختبارات الفرعية. أما طريقة التجزئة النصفية فقد حسبت لدرجات 70 عاملاً وتراوحت معاملات الثبات الناتجة عنها بين 0.5، 0.9 للنسب النكاه والاختبارات الفرعية. ولم يوضح الباحث كيف استخدم طريقة التجزئة النصفية لاختبائي رموز الأرقام وإعادة الأرقام بالرغم من عدم صلاحية هذه الطريقة لهما كما أشرنا.

## صدق المقياس

يفضل وكسلر في تناوله لصدق مقياسه أن يؤكد ' صدق المحقري ' . ومن رأيه أن الاختبارات الأحد عشر المتضمنة فيه تتفق مع تعريف الذكاء، وأن اختبارات مماثلة استخدمت في مقاييس الذكاء السابقة، وتبنت جدواها في الأبحاث الإكلينيكية. ومع ذلك فإنه يذكر بعض البيانات المحدودة عن الصدق التلازمي للمقياس. ومن ذلك المقارنة بين المجموعات للتطبيقية والمهنية المختلفة، بالإضافة إلى حساب معاملات الارتباط مع تقديرات الأداء على العمل والتدريبات المدرسية، وكانت أغلب الفروق بين المجموعات في الاتجاه المتوقع، ومن ذلك أن درجات الصال أعلى في الاختبارات العملية، وأن درجات الأفراد في المهون غير العملية أعلى في الاختبارات اللفظية. أما عن معاملات الارتباط بالمحكات المهنية والمدرسية فقد أعطى المقياس اللفظي قيماً أعلى من المقياس الكلي. أما المقياس الأدبي فقد أحلى أدنى قيم، وفي كل الأحوال كانت معاملات ارتباط المقياس اللفظي أدنى من معاملات مقياس ستانفورد بينيه. أما عن الصدق التكراري للمقياس فقد ثبت أن للمقياس يتكياً بدرجة مقبولة بالضعف العقلي كما تؤكد بحوث جيرتن وزملائه.

وقد حسبت معاملات الارتباط بين مقياس وكسلر وغيره من الاختبارات والمقاييس لمدة لقياس الذكاء، وكانت النتائج كما يلي:

- 1- معامل الارتباط بين المقياس ومقياس ستانفورد بينيه 0.80 وهو أعلى في المقياس اللفظي منه في المقياس الكلي والمقياس الأدبي.
- 2- معاملات الارتباط بين المقياس واختبارات الذكاء الجماعية تتراوح بين 0.40، 0.80.

3- معامل الارتباط بين المقياس الأدائي واختبار منبوسا لوحات الأشكال  
72.

4- معامل الارتباط بين المقياس الأدائي واختبار رافن للمصفوفات المتتمة  
0.70.

5- معامل الارتباط بين المقياس الأدائي واختبار بيتك للفهم الميكانيكي  
0.35.

### صدق التكوين الفرضي

أجريت عدة تحليلات عملية للاختبارات الفرعية بالتالي يتألف منها  
مقياس وكسلر. ومن أشهر هذه الدراسات بحوث كوهن على مصفوفات ارتباطية  
لمجموعات عمرية أربعة من عينة التتبعين هي: (من 18-19، ومن 25-35،  
ومن 45-54، ومن 60-75+). وقد وجد الباحث أن الاختبارات الفرعية جميعاً  
تتشارك في عامل عام واحد يسهم بحوالي 50% من التباين الكلي البطارية.  
وبالإضافة إلى ذلك استخراج ثلاثة عوامل طاقية هي:

1- عامل الفهم اللغوي وتشبع في اختبارات المفردات والمعلومات العامة  
والفهم العام والمنتزعات.

2- عامل لتنظيم الإدراكي وتشبع في رسوم المكعبات وتجميع الأشياء،  
وهو عبارة عن مركب من علمي السرعة الإدراكية والقصور البصري  
لمكاني الذين أكتنهما البحوث لعملية المرجحة وبخاصة بحوث  
نرستون.

3. عامل الذاكرة وتصبح أساساً في اختبارات الاستدلال الحسابي ومدى  
(إعادة) الأرقام، ويتضمن كلاً من الذاكرة الاستظهارية المباشرة للمواد  
للجديدة واستدعاء المواد التي سبق تعلمها.

ومن المثير أن نذكر أن نتائج هذا البحث لم تدعم الإجراء الذي  
يستخدم في تصنيف هذا المقياس أي تقسمه إلى مقياس لفظي ومقياس علمي  
لأن كلا منهما نسبة تكاء متفصلة. ورغم أن استخدام نسب التكاء الكلية يدعمه  
وجود عامل عام مشترك في جميع الاختبارات، فإن عامل الفهم اللفظي لم يظهر  
إلا في 4 اختبارات لفظية فقط من 6 اختبارات. أما عامل الذاكرة فقد وجد في  
الاختبارين اللفظيين الباقين بالإضافة إلى اختبارات فرعية أخرى من المقاييس  
عند الرشدنين الأكبر سناً. أما عامل التنظيم الإدراكي فكانت تشعباته الثلاثة  
لختبارين فقط من الاختبارات الأدائية الخمس، وأما الاختبارات الأدائية الثلاث  
الباقية فهي فيما يبدو لها عواملها الخاصة التي تشترك مع الاختبارات الأخرى  
المتضمنة في بطارية المقياس.

#### وصف مقياس وكسار تكاء الأطفال

يتكون مقياس وكسار تكاء الأطفال من الأسئلة السهلة والتي كان يتألف  
منها مقياس وكسار بقر الأصلي. ويقسم المقياس، كما هو الحال في مقياس  
الرشدنين، إلى قسمين: أحدهما لفظي والآخر علمي، والمجموع الكلي للاختبارات  
الفرعية 12 اختباراً منها اختباران احتياطيان أو إضافيان عند الضرورة، على  
النحو الآتي:

القسم اللفظي	القسم العملي (الآدائي)
1- المعلومات العامة	6- تكملة الصور
2- الفهم العام	7- ترتيب الصور
3- الاستدلال الحسابي	8- تصميم (رسم) المكعبات
4- التشابهات	9- تجميع الأكتاف
5- المفردات	10- للترميز
(إصادة الأرقام)	(المتاهات)

و قد اعتبر وكسلر الاختبارات الاحتياطية هي تلك التي ارتبطت لدى الارتباطات ببقية المقياس، وعلى ذلك فإن اختبار مدى (إصادة الأرقام في المقياس اللفظي واختبار المتاهات في المقياس العملي من هذا النوع. ويتطلب اختبار للترميز مع اختبار رموز الأرقام في مقياس للراشدين فيما عدا غلبة الأمثلة السهلة عليه. أما الاختبار للوحيد الذي ليس له نظير مقياس في الراشدين فهو اختبار المتاهات، ويتكون 8 متاهات ورافة متزايد الصعوبة، وتصحح في ضوء الزمن وعدد الأخطاء. وفي رأي وكسلر أن اختباري الترميز والمتاهات يمكن أن يحل أحدهما محل الآخر، ويعتمد ذلك على لقرار الذي يتخذه الفاحص. ويتطلب اختبار الترميز زمناً أقل من اختبار المتاهات، وقد يفسله بعض الفاحصين لذلك.

ولا يختلف مقياس الأطفال في إجراءات تصحيحه من مقياس الراشدين إلا اختلافات طفيفة. وتتخصص في تحويل الدرجات الخام في كل اختبار فرعي إلى درجات معيارية اعتدالية في إطار المجموعات العمرية التي ينتمي إليها المفحوص. وقد أعد وكسلر جداول لهذه الدرجات الموزونة لغات عربية



مقدارها أربعة شهور ما بين سن 5 سنوات حتى سن 15 سنة. وتشبه الموزونة هنا درجات مقياس قراندين في أن متوسطها 10 وانحرافها المعياري 3، ثم تجمع الدرجات الموزونة وتحويل إلى نسب ذكاء انحرافية متوسطها 100 وانحرافها المعياري 15 سواء للمقياس اللفظي أو للمقياس الأنلي أو للمقياس الكلي. وكذلك اهتم وكسلر باقتراح بعض الطرق الإحصائية لحساب الأعمار العقلية المقابلة لدرجات مقياس الأطفال. ورغم أن هذه الأعمار العقلية ليست من خصائص نسب الذكاء الانحرافية التي تتكف منها معايير المقياس إلا أنه إذا كان في ذلك يواجه مطلباً عالياً ملحاً نتيجة لشروح مفهوم العمر العقلي بين خاصة الناس وعملتهم بالنسبة لمبدأ قياس الذكاء. وقد أعدت بالفعل جداول لتحويل الدرجات الخام في المقياس إلى أعمار عقلية.

وتكلفت عينة تقنين مقياس وكسلر للأطفال من 100 طفل و 100 طفلة في الأعمار من 5 سنوات إلى 15 سنة، وكان المجموع الكلي للعينة 2200 مفحوصاً كلهم من تلاميذ المدارس باستثناء 55 حالة من حالة الضعف العقلي اختبرت من المؤسسات.

### ثبات المقياس

حصيت معاملات الثبات بطريقة التجزئة النصفية لكل لغتبار فرعي في مقياس وكسلر للأطفال ولكل من المقياس اللفظي أو المقياس العملي. والمقياس الكلي، لعينات  $7\frac{1}{2}$  سنة،  $10\frac{1}{2}$  سنة، 13  $\frac{1}{2}$  سنة، واحتوت كل عينة عسرية على 200 مفحوص، ولم تستخدم طريقة التجزئة النصفية مع اختباري الترميز ومدى الأرقام لعدم ملائمتها (لأسباب وضاحتها آنفاً)، ولذلك حسب الثبات لها بطريقة الصور المتكافئة، وكانت معاملات الثبات للمقياس الكلي للمجموعات

المصرية الثلاثة 0.92، 0.95، 0.94، وعلى التوالي، والمقياس اللفظي: 0.88، 0.99، 0.96 على التوالي، والمقياس الأدبي: 0.85، 0.89، 0.90 على التوالي. وهي معاملات مرتفعة وتدل على ثبات المقياس.

لما معاملات ثبات الاختبارات الفرعية فكانت منخفضة بصفة عامة فيما عدا اختبار للمفردات كما أن هذه المعاملات كانت أكثر انخفاضاً في حالة الأطفال الصغار منها في حالة الأطفال الكبار.

#### صدق المقياس

لا تتضمن كراسات تعليمات مقياس وكسور للأطفال بيانات عن صدقه، ولو أن بعض الشواهد تدل على ذلك، وخاصة ما لوحظ من أنه من حالة واحدة من بين 55 حالة ضعف عقلي تضمنها هيئة التقنين وصل مستوى ذكائها في المقياس اللفظي إلى نسبة ذكاء انحرافية 70، وكان متوسط نسب ذكاء هذه المجموعة 57. كما أثبت عدد من الباحثين الصدق التلازمي للمقياس باستخدام محكات اختبارات التحصيل الأكاديمي، وبلغت معاملات الارتباط حوالي 0.60، وكانت معاملات ارتباط المقياس اللفظي بهذه المحكات بالطبع أعلى من المقياس العملي.

لما بحوث التحليل العامل التي أجريت على المقياس بهدف تحقيق صدق التكوين الفرضي والتي يلخصها ليدل أما عام 1960 فقد استخرجت عوامل تشبه ما وجد في تحليل مقياس الراشدين وهي: العمل العام، والقيم اللفظي، والعمل المكاني الإدراكي، وعامل الذكاء. ولا يوجد إلا قدر قليل من

التطابق بين عامل الفهم اللفظي والاختبارات الفرعية. المقياس اللفظي، والعامل  
المكاني الإدراكي والاختبارات الفرعية في المقياس الأكاديمي.

### تكوين المقياس

يمكن القول أن مقياس وكسلر للأطفال مقياس جيد لذكاء الأطفال ويتفق  
في الخصائص مع كثير من الاختبارات الفرعية. ومع ذلك فمن رأي استناري  
أن الأساس المنطقي لهذا المقياس ظاهري التناقض. فقد أشرنا إلى أن مقياس  
وكسلر بلفظي الأصلي لمقابلة الحاجة إلى اختبار الراشدين التي لا يوفرها مقياس  
متلفرد ببنوه أو أي مقياس آخر يعتمد على مجرد (رفع مستوى) مقاييس  
الأطفال. وعندما حقق وكسلر هذا الهدف سمى إلى إعداد مقاييس للأطفال تعتمد  
في جوهرها على (خفض مستوى) مقاييس الراشدين، بل أنه صار بالتناقض إلى  
نهائيه حينما أعد مقاييساً للأطفال ما قبل المدرسة.

### مقياس وكسلر لذكاء الأطفال ما قبل المدرسة

ظهر مقياس وكسلر لذكاء الأطفال ما قبل المدرسة عام 1967، ويهدف  
إلى قياس ذكاء الأطفال من سن 4 سنوات إلى  $6\frac{1}{2}$  سنة، ويتألف من 11  
اختباراً فرعياً يستخدم منها 10 اختبارات فقط في تحديد نسبة الذكاء، ومنها نجد  
8 اختبارات عبارة عن " خفض لمستوى " الاختبارات المتداخلة لها في مقياس  
الأطفال أنفسهم للذكر، أما الاختبارات الثلاثة الباقية فقد أعدت خصيصاً لهذا  
المقياس. وتصنف الاختبارات الفرعية في هذا المقياس، كما هو الحال في  
مقياسي Wisc, Weis إلى قسمين: لفظي وعلمي، على النحو الآتي (وقد  
وضعنا علامة \* أمام الاختبارات الجديدة).

المقياس اللفظي	المقياس الأدائي (العملي)
المعلومات	* بيت الحيوان
المفردات	تكلمة للصور
الحساب	المتاهات
* المتشابهات	* الرسوم الهندسية
الفهم	رسم المكعبات
* الجمل (اختبار إنشائي)	

ويوصى وكملر باستخدام الاختبارات اللفظية والعملية بالتناوب لتوفير التنوع والاحتفاظ بميل الطفل وتعامله. ويتطلب المقياس في تطبيقه ما بين 50 دقيقة و 75 دقيقة في جلسة واحدة أو جلستين اختباريتين.

ولقد حل اختبار الجمل في هذا المقياس محل اختبار مدى (إعادة) الأرقام في مقياس Wisc، وفيه يكرر الطفل كل جملة مباشرة بعد عرضها عليه شفويًا. وهذا الاختبار يمكن أن يستخدم كبديل لأي اختبار لفظي آخر، أو يطبق كاختبار إنشائي لتزويد الفلخص بمعلومات أكثر عن الطفل، وفي كلتا الحالتين لا يستخدم في حساب نسبة الذكاء. أما اختبار بيت الحيوان فيشبه اختبار رموز الأرقام في مقياس Wais واختبار الترميز في مقياس Wisc. ويتضمن اختبار مفتاحاً يعطى للطفل على لوحة يتألف من صور لكاتب ودجاجة وسكة وقطة لكل منها بيت (المطواشي) ملون بلون مختلف موضوع تحت الصورة، وعلى الطفل أن يضع الأسطوانة (البيت) ذات اللون المناسب في فتحة تحت صورة الحيوان على اللوحة. وتتحدد درجة لطفل في ضوء الزمن وعدد الأخطاء وعدد الاستجابات المحذوفة. أما اختبار \* الرسوم الهندسية \* فيطلب من الطفل محاكاة 10 رسوم بسيطة بقلم ملون.



# اختبارات القدرة المتخصصة





## اختبارات القدرة المتخصصة

أشرنا في الفصلين السابقين إلى أن اختبارات الذكاء العام لا تشمل فيما تقيسه ميدان " القدرات العقلية "، وقد ظهر هذا للنقص حتى قبل الاهتمام بهطاريات الاستعدادات للمتحدة، وتمثل ذلك في ظهور الاختبارات التي تقيس ما كان يسمى " الاستعدادات الخاصة " وأشهرها وأكثرها تذكيراً اختبارات " الاستعداد الميكانيكي "، ثم تزايدت الحاجة إلى هذه الاختبارات التي تقيس ما يمكن أن نسميه " القدرات المتخصصة " بزيادة الاهتمام في علم النفس التطبيقي بمشكلات التوجيه والانتقاء والتوزيع والإرشاد النفسي، فكثر عددها وشملت مجالات أكثر تنوعاً ومن ذلك ميادين النشاط الموسيقي والفني والمهني والكتابي، بالإضافة إلى النشاط الميكانيكي الذي أشرنا إليه.

وقبل أن نتناول نماذج من اختبارات قدرات المتخصصة بشيء من التفصيل لا بد أن نوضح مكانة هذه الفئة من الاختبارات في النظرية الحديثة للقدرات العقلية ولنجاز نقول أن هذه الاختبارات ظهرت في وقت كان الاهتمام كله مركزاً على ما يسمى " اختبارات الذكاء العام " والتي أشرنا إليها في الفصل الرابع، بهدف أن تكون مكملة لمفهوم " نسبة الذكاء " في وصف النشاط العقلي المفصوص، إلا أنه مع التقدم في مناهج البحث وفي نظرية قدرات العقلية، وخاصة بعد شيوع استخدام منهج التحليل العامل والنماذج المؤسسة عليه، أصبح من الشائع اعتبار مفهوم " الذكاء " ذاته يتكف من عدد كبير من " القدرات العقلية " يمثل في أحد النماذج المعاصرة ونقصد به نموذج جيلفورد إلى أكثر من 120 قدرة وزاد نتيجة لذلك الاهتمام بهطاريات الاستعدادات المتخصصة. بل إن



بعض الاختبارات التقليدية \* للاستعدادات الخاصة \* أصبحت من مكونات بعض بطاريات الاستعدادات كما أشرنا في الفصل الخامس.

والسؤال الذي تطرحه امتنازى هو: ما هو موضع هذه الاختبارات في الوقت الحاضر ؟

الواقع أن لهذه الاختبارات وظيفتين هامين في التقويم النفسي المعاصر هما:

1- توجد بعض مجالات السلوك الإنساني لا تتوافر لها مقاييس جيدة في بطاريات الاستعدادات المتعددة ومن ذلك الإدراك البصري والسمعي والمهارة الحركية والمواهب الموسيقية والفنية والسلوك الابتكاري. وقد يكون السبب في ذلك أن المواقف التي تتطلب استخدام هذه الاختبارات على درجة كبيرة من النوعية بحيث لا تكرر إجرائها في بطاريات الاستعدادات المتحدة. ومعنى ذلك أن هذه الاختبارات تستل فئة متميزة من الاختبارات العقلية حتى تتوافر بطاريات \* شاملة \* لقدرات الإنسان. وهنا ما ننتبأ حدوثه في المستقبل قريبا لعلم القياس النفسي.

2- تستخدم اختبارات القدرات المتخصصة حتى في الأحوال التي توجد نماذج لها في بطاريات الاستعدادات المتحدة، ومن ذلك اختبارات القدرات الميكانيكية أو الكتابية. والسبب في ذلك أن بعض اختبارات القدرات المتخصصة تتوافر لها بيانات ثقلين وافية، بالإضافة إلى المعونة في استخدام الاختبارات، وشمول الميدان الذي تقيسه. بدلاً من استخدام اختبار واحد القدرة الميكانيكية في إحدى البطاريات المتحدة الاختبارات، يمكن للباحث أن يطبق بطارية لاختبارات القدرة.

ويؤاخر في الوقت الحاضر عدد هائل من القدرات التي تقيس \* القدرات المتخصصة \* ابتداء من المستوى الحاسي (السمع والبصر) حتى مستوى التفكير الابتكاري والنشاط المهني. ولا يتسع المقام لتناول هذه الاختبارات بأي نوع من العرض الشامل ولذلك سوف تقتصر على طائفة من هذه القدرات وما يقيسها من اختبارات.

### لختبار القدرات الحركية

ظهرت الاختبارات التي تقيس خصائص الامتجابات الحركية كالسرعة والفتور منذ فترة طويلة، ويهتم معظمها بالمهارة اليدوية وقليل منها يهتم بحركات الساق والقدم والتي تتطلبها بعض المهن والأعمال. وقيس بعضها مزيجاً من الاستعدادات الحركية والإبرلية والمكانية والميكانيكية. وفي أغلب الأحوال فإن اختبارات القدرات الحركية تتطلب أجهزة في قياسها رغم ظهور بعض اختبارات الورقة والقلم في أغراض التقويم النفسي الجماعي. وقد أشرنا في الفصل الخامس إلى بطاريني فلانجان والاستعدادات العامة اللتين تضمنتا بعض الاختبارات الحركية بالجماعية إلا أن الدراسات التي قام بها فاشمان وملتون تؤكد عدم وجود علاقة بين اختبارات الورقة والقلم واختبارات الأجهزة التي تقيس نفس القدرات الحركية، وفيما يلي أمثلة للأجهزة التي تستخدم في قياس للنشاط الحركي:

- 1- مقياس القوة العضلية (الدينامومتر) Dynamometer وتستخدم في قياس قوة قبضة اليد.
- 2- جهاز قياس زمن الرجوع.

3- اختبار Grawford Small Parts Dexterity ويقيس المهارة اليدوية البسيطة.

4- اختبار Purdue Pegboard ويقيس النشاط الحركي الدقيق والنشاط الحركي الخفيف.

5- اختبار Stomberg Dexterity ويقيس حركات اليد والذراع.

### الختبارات القدرات الميكانيكية

إن ما يسمى القدرة الميكانيكية ليس قدرة موحدة لا تقبل التحليل، وإنما هي في الواقع مركب من القدرات الحسية والإدراكية والحركية، بالإضافة إلى قدرات إدراك العلاقات المكانية، واكتساب المعلومات الميكانيكية، وفهم العلاقات الميكانيكية.

وإنما فإن الاختبارات التي تقيس القدرة الميكانيكية إنما تهتم بمستوى من الأداء أعلى مما تتضمنه الاختبارات الحسية والإدراكية (لتمييز السعي والبصري والحركي والفتقر الحسلي والمهارة اليدوية).

ويعد اختباراً للتجميع Assembly الذي أعده ستوكويست عام 1923 أول اختبار ظهر في هذا الميدان يقيس بقدرة المفحوص على تركيب أجزاء الأجهزة الميكانيكية مثل جرس الدراجة، أو قفل، أو مصيدة فئران، ويتألف من 3 سلاسل تستخدم مع مستوى عصري يمتد بين الطفولة والرشد.

وقد حُل هذا الاختبار في جامعة مينيسوتا عام 1930 وأصبح يطلق عليه اختبار مينيسوتا للتجميع الميكانيكي وأضيفت إليه أجهزة جديدة. ويصحح هذا الاختبار كسابقه في ضوء معدل الاستجابة ونقته. وقد أُنشئت الدراسات التي

أجريت عليه أنه على درجة كافية من التنبؤ بدجاج طلاب لتطعيم الفتي المبتدئين في أصل الورشة، كما ارتباط بالدجاج في بعض المهن الميكانيكية.

وفي عام 1930 ظهر اختبار آخر في جامعة ميليسوتا هو اختبار ميليسوتا للعلاقات المكانية ويتألف من سلسلة من 4 لوحات خشبية وتكون كل منها من 57 قطعة خشبية مختلفة الأشكال، وبمضنها غير مألوف ويطلب من المفحوص أن يضع هذه القطع لساكنها في فتحات اللوحة الخشبية، وفي رأي كثير من النقاد أن هذا الاختبار يقيس المعرفة والقدرة في الاستجابة لتفاصيل العلاقات المكانية، والقدرة المفحوص على التعامل مع الأشياء والمواد المصنوعة. وبعبارة أخرى فإنه لا يقيس القدرة على حل المشكلات ذات الطبيعة الميكانيكية، كما لا يصلح لقياس القدرة على معالجة الأشياء الصغيرة بدقة.

وفي عام 1948 ظهر اختبار ميليسوتا للوحات الأشكال الورقية يعرض على المفحوص على ورقة مطبوعة نفس المشكلات التي كانت تعرض عليه في لوحات الأشكال الخشبية. وفي كل سؤال من أسئلة هذا الاختبار يحرص على المفحوص جزءان أو أكثر من شكل هندسي بحيث لو جمعت معاً بطريقة صحيحة فإنها تكون الشكل الكلي للكل. وعلى المفحوص تحديد الشكل الذي يمكن أن يجمع مع الأجزاء الأخرى من بين 5 بدائل اختيارية. وهكذا يقيس هذا الاختبار القدرة المكانية كما تحدث في البحوث العملية للقرارات العقلية. وتؤكد نتائج البحوث التي أجريت على هذا الاختبار أنه يرتبط ارتباطاً دالاً ببلوغ الأداء الميكانيكي، والدجاج في الرسم الهندسي والهندسة الوصفية. واستخدمت في حساب صدقه محكات مختلفة مثل درجات مقررات الهندسة والمواد الفيزياء والورش الميكانيكية وتقديرات المشرفين ومقياس الكفاءة الإنتاجية.

وتوجد فئة أخرى من الاختبارات القادرة للميكانيكية بقياس المعلومات للميكانيكية، والاستدلال الميكانيكي والفهم الميكانيكي، وهذه جميعاً تتطلب من المفحوص لفئة بالآلات للشائعة وبالعمليات الميكانيكية وبالعمليات الميكانيكية، إلا أنها لا تقتصر في المفحوص تولف معلومات فنية أكثر مما يمكن أن توفره الحياة اليومية في المجتمعات الصناعية. ومن أشهر هذه الاختبارات تلك التي أعدها بينت وزملاؤه والتي تسمى اختبارات الفهم الميكانيكي وتعرض على المفحوص مشكلات ميكانيكية مسورة وعدة حلول لكل منها، وعليه أن يختار الحل الصحيح. ولهذه الاختبارات ثلاثة مستويات من الصعوبة، وتهدف إلى قياس القدرة على فهم المبادئ الفيزيائية والميكانيكية في مواقف بسيطة نسبياً.

ومن الاختبارات الأحدث نسبياً في قياس القدرة للميكانيكية اختبار الدافعية الميكانيكية المسمى Mellanbruch Mechanical Motivation وهو اختبار مصور يتطلب من المفحوص التعرف على أشياء معينة شائعة في المجتمعات الصناعية المتقدمة. والاختبار الثاني هو SRA Mechanical Aptitudes Test ويتألف من 3 اختبارات فرعية تقيس:

- أ- المعلومات الميكانيكية.
- ب- إدراك الأشكال والتصور البصري المكاني.
- ج- حل المشكلات الحسابية الخاصة بالأحصال الميكانيكية (في صورة جدول ورسوم بيانية).

ويجب أن نؤكد أن هذه الاختبارات التي تقيس للفهم الميكانيكي تتطلب تحديلاً شاملاً عند محاولة استخدامها في البيئة المصرية، وذلك لأنها مضبوطة تشبهاً كبيراً بالعناصر الثقافية السائدة في المجتمعات التي صنعت فيها.

## الختبار للقدرة الكتابية

للقدرة الكتابية - كالقدرة الميكانيكية - ليست قدرة بسيطة وإنما هي على درجة من التعقد والتركيب. ورغم أن بعض بطاريات الاستعدادات للمتحدة التي تناولناها في الفصل الخامس - وخاصة بطارية الاستعدادات الفارقة وبطارية الاستعدادات العامة - تضمنت اختبارات لدرجة لقياس القدرة الكتابية، فإن تراث التعويم للنسبي يتضمن طائفة من الاختبارات التي تنبئ هذه القدرة على وجه الخصوص، بل توجد بطاريات لقياس القدرة الكتابية، وعموماً نستطيع القول أن اختبارات هذه القدرة تجمعها خصائص مشتركة يوضحها الجدول الآتي الذي يلخص محتوى 6 بطاريات للقدرة الكتابية ولختباراتها الفرعية. ومنه يتضح أن البطاريات الشاملة تسمى إلى قبل عدد من العيول النوعية، بينما تسمى البطاريات الأقل شمولاً إلى قبل السرعة والفتة الإدراكيين والتي تعد أحد جوانب القدرة الكتابية وليست كلها. وقد يكون من العوامل الكامنة وراء قصر القدرة الكتابية على السرعة الإدراكية ما يؤكد نتائج البحوث من أهمية القدرات الإدارية بوجه عام فيها. إلا أن " القدرة الكتابية " تشمل أنواعاً متنوعة من الأعمال تختلف في مدى القدرة ومستواها، ولذلك فإن مقاييسها تتضمن عيولاً عقلية أخرى أساسية، بالإضافة إلى بعض القدرات الحسية والمهارات اليدوية عند الضرورة.

الاختبارات الفرعية	بطارية
<p>للخط: السرعة والدقة</p> <p>للمراجعة: السرعة والدقة</p> <p>لحساب البسيط</p> <p>للمرعة والدقة في الأداء الحركي</p> <p>معرفة المصطلحات التجارية</p> <p>للبسيطة</p> <p>ترتيب الصور</p> <p>للتصنيف: السرعة والدقة</p> <p>لترتيب الأبجدي</p>	<p>بطارية ديترويت</p> <p>Detroit</p>
<p>للمزاجعة: اكتشاف الأخطاء في</p> <p>الأسماء والأعداد ، لترتيب الأبجدي</p> <p>لحساب: الجمع البسيط</p> <p>لحساب: تحديد أخطاء الجمع</p> <p>لمسائل الحسابية البسيطة</p> <p>لتهجي</p> <p>فهم القراءة</p> <p>معرفة معاني الكلمات</p> <p>الاستخدام اللغوي: القواعد النحوية</p>	<p>بطارية الاستعدادات الكتابية العامة</p> <p>General Clerical</p>

الاختبارات الفرعية	البطارية
مقارنة الأعداد مقارنة الأسماء	بطارية مينيسوتا
التهجي الحساب المراجعة: السرعة معرفة معاني الكلمات النقل والنسخ: الدقة الاستدلال	بطارية بورد Purdue
العمليات العددية معاني الكلمات التصنيف	بطارية التوظيف Short Employment
المهارات اللفظية المهارات العددية التعليمات المكتوبة المراجعة: السرعة التصنيف الترتيب الأبجدي	بطارية تورس Turse

بد في المكتبة العربية بطاريتان لقياس القدرة الكتابية هما:

(1) اختبارات القدرة الكتابية للدكتور محمد عبد السلام أحمد

ويتألف من اختبارين هما:



- اختبار تصنيف الأعداد.

- اختبار تصنيف الأسماء.

اعتماداً على تطويل للأعمال الكتابية قام به الدكتور محمد عبد السلام أحمد بمعاونة ديوان الموظفين عام 1953. وكان الاختبار الأول على صيغة حجمها 822 مفهوماً من الجنسين من الحاصلين على شهادة الثانوية للعلماء أو دبلوم للتجارة الثانوية تتراوح أعمارهم بين 18، 28 سنة. وكان الاختبار الثاني على صيغة عددها 880 فرداً من الجنسين تتراوح أعمارهم أيضاً بين 18، 28 سنة من حملة الثانوية العامة المتعلمين لثقل الوظائف الكتابية عن طريق إعلانات ديوان الموظفين في ذلك الوقت. وأحدث للاختبارين معياراً مئينية وثاقية وحسب لهما معامل ثبات بطريقة التجزئة للنصفية فبلغ 0.97 والواقع أن استخدام طريقة التجزئة للنصفية في حساب اختبارات القدرة الكتابية - التي تعتمد اعتماداً كلياً على السرعة - يثير كثيراً من المشكلات أضربنا إليها في الفصول السابقة، إلا إذا كان المؤلف لجأ إلى تجزئة زمن الاختبار، وهذا ما لم توضحه كراسة تعليمات هذه الاختبارات. أما عن صدق الاختبارين فقد حسب مؤلفهما معاملات الارتباط بينهما وبين اختبارات المهن الكتابية وللذكاء الاجتماعي فلو حظ أن الاختبارين يشتركان مع الاختبارات التي تقاس الجوانب الأكبية؛ السرعة، الدقة، سهولة استخدام اللغة، الألفة بالأعداد، للذاكرة المباشرة، المهارة اليدوية.

(2) اختبارات المهن الكتابية من إعداد الدكتور محمد عماد الدين إسماعيل

والدكتور سيد عبد الحميد مرسى:

وتتكون مكونات ثلاث هي:

- القدرة للحذبة.

- السرعة والدقة،
- الاستدلال اللغوي.

وقد حسب ثبات الاختبارات للثلاث بطريقة التجزئة النصفية فبلغت 0.86، 0.84، 0.85 على التوالي. ومرة أخرى نشر مسألة استخدام الطريقة الملائمة لحساب ثبات الاختبارات التي تعتمد على السرعة. واستخدم في حساب صدق الاختبارات تقديرات مادة المحاسبة والترجمة وتقديرات المدرسين والمعلمين وتلويحت معاملات الارتباط بين 0.42، 0.54 وحسبت المعايير المئينية للاختبارات.

### الختبارات القدرية الموسيقية

وجد في الفترات الحاضر عدد قليل من الاختبارات التي تصلح لقياس القدرة الموسيقية؛ قد يكون أهمها ظهور أو أكثرها أهمية لاختبارات سيشور التي تسمى: Seashore Measures of Musical Talent.

وقد ظهرت في صورتها الأصلية عام 1919، واستمرت شائعة الاستخدام في الولايات المتحدة دون تعديل حتى عام 1939 حيث عكفت في صورتها التي تستخدم حتى الآن. وقد كانت هذه الاختبارات نتيجة للبحوث الواسعة النطاق في ميدان سيكولوجية الموسيقى والتي أجريت بجامعة يوتا تحت إشراف كارل سيشور. وتتألف هذه البطارية في صورتها الحديثة من 6 اختبارات هي تمييز الأصوات، شدة الصوت، تذكر الإيقاعات، الزمن، نوعية الصوت، تذكر الأكالن.

ويتكلف اختبار تمييز الأصوات من 50 زوجاً من النغمات، وعلى المفحوص أن يحدد بالنسبة لكل زوج ما إذا كانت النغمة الثانية أهد أم أعلظ من النغمة الأولى، وفي اختبار شدة الصوت يطلب منه أن يحدد بالنسبة لكل زوج من 30 زوجاً من النغمات ما إذا كانت النغمة الثانية أشد أم أضعف من الأولى. ويطلب منه في اختبار تكرر الإيقاعات أن يحدد ما إذا كان النمطان الإيقاعيان اللذان يتكلف منهما كل سؤال (من 30 سؤالاً) متشابهين أو مختلفين. وفي اختبار الزمن يحدد المفحوص ما إذا كانت النغمة الثانية أطول أم أقصر من النغمة الأولى (50 سؤالاً) وفي اختبار نوعية الصوت يقوم للمفحوص بالحكم على النغمة الثانية بأنها متشابهة أو مختلفة عن النغمة من حيث نوعية الصوت (البناء الهارموني)، ويتكون هذا الاختبار من 50 سؤالاً. أما الاختبار الأخير، وهو تذكر الأغنان، فيتكون من 30 زوجاً من التماذج اللفظية، وفي كل زوج توجد نغمة واحدة مختلفة في النموذج الثاني، وعلى المفحوص أن يحدد أي النغمات اختلفت ويحدد ترتيبها الرقمي.

وتصلح هذه الاختبارات للتطبيق على المفحوصين ابتداء من الصف الرابع الابتدائي حتى أعلى المستويات التي تتطلب إعداداً موسيقياً رفيعاً للمراهقين والراشدين.

وقد ظهرت هذه الاختبارات في أسطها الأجنبي مسجلة على أسطوانات، وقد أعدتها باللغة العربية الدكتورة أمل أحمد مختار صديق مسجلة على شرائط، وأعدت تعليمات تطبيقها، كما صممت كراسة إجابة مختلفة عن الطبعة الأجنبية. وحسبت معاملات الثبات بطريقة التجزئة النصفية وتصحيح سبيرمان درلون على عينة من 163 طالباً وطالبة من حملة الثانوية العامة للمتقدمين

الاتحاق بالمعهد العالي للتربية الموسيقية وتراوحت معاملات الثبات بين 58 واختبار الزمن 88 واختبار تذكر الأكلان. أما عن صدق هذه الاختبارات فقد حسبت الباحثة معاملات الارتباط بين الاختبارات القرعية التي تتألف منها البطارية لعينات من طلاب دور المطين والمطبات وطلاب المعاهد الموسيقية العليا وتلاميذ المدارس الإعدادية العامة، وتلاميذ المعهد القومي للموسيقى، كما حسبت معاملات الارتباط بين اختبارات سيشور واختبارات القدرات الموسيقية للأطفال (بناتي) لعبة تلاميذ المرحلة الإعدادية العامة، وتلاميذ المعهد القومي للموسيقى، وكذلك معاملات ارتباطها باختبارات ونج للذكاء الموسيقي، ولوريجون للتفريق الموسيقي، وفارم للتكوين الموسيقي، وكامل للذكاء العام، واختبار الذكاء المصور للدكتور أحمد زكي صالح، وتتراوح في كرامة بالتحليلات بيانات كافية عن المعالم الميكرومتريّة لهذه البطارية. أما عن معايير هذه البطارية فقد حسبت المنحنيات العامة بكل اختبار فرعي ولكل عينة فرعية من عينة للتقنين.

ويوجد اختبار آخر للقدرّة الموسيقية ظهر عام 1963 أعده في إنجلترا أرنولد بنتلي خصيصاً لأطفال المرحلة الابتدائية، (من 7-12 سنة) ويتيسر 4 قدرات هي: تمييز الأصوات، تذكر النغمات، تحليل النغافات، تذكر الإيقاعات. ويمكن تطبيق هذه الاختبارات على مستويات عمرية أعلى مع مراعاة معايير هذه الأصار عند تفسير الدرجات.

ويتكون لاختبار تمييز الأصوات من 20 زوجاً من الأصوات ' للثقة ' يستمع إليها المفحوص ثم يطلب منه أن يحدد ما إذا كان الصوت الثاني في كل زوج متطابقاً مع الصوت الأول أو مختلفاً عنه. وفي حالة الاختلاف على

المفحوص أن يحدد ما إذا كان أظلى (أي أحد) أو أسفل (أي أظلم) من الصوت الأول. أما اختبار تذكر اللغات فيكون من 10 لسئلة ويطلب كل منها مقارنة ثنائية بين جملتين لطيفين قصيرتين، ويتكون كل منها من 5 لغات، ويطلب من المفحوص أن يحدد ما إذا كانت الجملة الثانية متشابهة مع الأولى أو مختلفة عنها وإذا حدث اختلاف عليه أن يحدد ترتيب اللغمة المختلفة. وفي اختبار تحليل التلغات يطلب من المفحوص تحديد عدد اللغات التي يتكون منها كل تليف (يتكون الاختبار من 20 تليفاً). أما اختبار تذكر الإيحاءات فيتكون من 10 مغردات ويطلب كل منها مقارنة ثنائية بين جملتين إيحائيتين، وتتكون كل واحدة من 4 وحدات إيحائية أو ما يعادلها من النماذج الإيحائية المختلفة، ويطلب من المفحوص تحديد ما إذا كانت الجملة الثانية لي كل سؤال متشابهة مع الجملة الأولى أو مختلفة عنها، وفي حالة الاختلاف يكون عليه تحديد رقم الوحدة

#### النتائج

وقد أحدثت هذه الاختبارات باللغة العربية للذكورة آمال أحمد مختار صادق عام 1972، وظهرت اختبارات القدرات الموسيقية للأطفال مسجلة على شرائط (بدلاً من الاسطوانات في الأصل الأجنبي) وأصبحت لها كراسة إجابة خاصة، كما فصلت تعليمات تطبيقها. وقد استخدم المؤلف الألفي في حساب ثبات هذه الاختبارات طريقة إعادة الاختبار، وبلغ معامل الارتباط 0.84. واستخدمت الباحثة المصرية هذه الاختبارات في دراستين عاملتين للقراءة الموسيقية إحداهما على عينة من تلاميذ المرحلة الإعدادية، والأخرى على عينة تتألف منها في مصر الرمزي من تلاميذ المعهد القومي للموسيقى (الكونسيرفاتور). عددها 74 تلميذاً.

ولصاحب صدق الاختبارات حصلت معاملات الارتباط بين الاختبارات  
الفرعية، ومعاملات ارتباطها باختبارات ميسور التي أشرنا إليها، وباختبارات  
كلل للنكاء العام. وكذلك حصلت المعايير المعنوية لمؤشرين من التلاميذ: إحداهما  
مجموعة 9-12 سنة والأخرى مجموعة 12-15 سنة.

ويوجد اختبار آخر ظهر في إنجلترا عام 1941 ثم ظهرت طبعته  
الجديدة المعدلة عام 1958 ثم عام 1962 وهو اختبار ونج للنكاء الموسيقي.

ويصلح للتطبيق ابتداء من سن 8 سنوات، ويختلف عن اختبار ميسور  
في أنه لا يؤكد للجانب الفرعي التحليلي كما يشمل في عناصر الإحساس البسيط  
المساعد في اختبارات ميسور، ولذلك فإنه يعتمد على مستوى موسيقي واقعي  
حيث تستخدم موسيقى البيلو في كل اختبار من الاختبارات السبع التي تتألف  
منها بطارية ونج وهي: نطيل لتكلفت، تمييز الأصوات، ذاكرة الأصوات،  
لتوافق الموسيقي (الهارموني)، الشدة، الإيقاع، الجدل الموسيقية. وتتطلب  
الاختبارات الثلاث الأولى تمييزاً حسيّاً على مستوى أكثر تحضراً من لاختبارات  
ميسور، أما في الاختبارات الأربع الأخرى فإن المفحوص يقارن النوعية  
الجمالية في مقطوعتين موسيقيتين، ومعنى ذلك أن بطارية ونج تؤكد جانب  
التنسيق الموسيقي. ويقترح ونج استخدام الاختبارات الثلاث الأولى كبطارية  
مختصرة مع الأطفال الصغار أو كأداة تصفية مبدئية.

#### الاختبارات المقررة في الفنون البصرية

توجد في الوقت الحاضر طائفة من الاختبارات التي تقبل بعض  
جوانب القدرة فيمدان الفنون البصرية، ومنها اختبار Graves Design

Judgment الذي يهدف إلى قياس قدرة المفحوص على إدراك المبادئ الأساسية لتعمل الجمالي والاستجابة لها وهي مبادئ الوحدة والسيولة والتكثيف والتوازن والاستمرار والتكيف والتناسب والإيقاع. وقد تحقق ذلك بمرض 90 ثنائية أو ثلاثية من الرسوم المجردة، في كل منها واحد من الرسوم يتفق مع المبادئ الثنائي هذه بينما يختلف أحد هذه المبادئ أو بعضها في الرسم الآخر أو الرسوم الأخرى. ويطلب من المفحوص أن يفضل بين رسوم الثنائية أو الثلاثية ويختار أحدها على أساس التفضيل الفني. ولم يتضمن هذا الاختبار نماذج من أعمال الفن، لأن الحكم على الصور واللوحات يتأثر بالخبرة السابقة وسمات الشخصية والانفعالات ومفهوم المفحوص ص الفن، ولذلك فإن الاختبار يؤكد محض الاختيار للجمالي، أو ما يسمى (الحساسية الجمالية). وبالإضافة إلى ذلك فإن الاختبار يقيس للصلاحيات الإدراكية المعقدة وليس العناصر البسيطة البسيطة التي تتدخل في الحكم الجمالي، وفي هذا يتشابه مع اختبارات ورج في الموسيقى، ومع المبادئ المتضمنة في قياس العامل في الذكاء.

والاختبار الثنائي في هذا المجال هو اختبار ملير للحكم الفني. ويهدف إلى قياس الحكم الجمالي بطريقة (كافية). ويختلف عن اختبار جريفي في أنه يتكلف من 100 زوج من اللوحات الفنية مرسومة بالأبيض والأسود، ويسل أحد عناصر ثنائية صلاً فنياً ممتازاً، والآخر أدخل عليه بعض التغيير والتعديل في أحد الجوانب الفنية الهامة تملأ كل جودة من الحصر الأول. ويحدد للمفحوص أي الجوانب قد تغير (الشكل أو الزوايا مثلاً) إلى أنه لا يعلم أي اللوحين هو اللوحة الأصلية، ويطلب من المفحوص أن يحدد تفضيله لأي منهما.

ويرى ملير أن هذا الاختبار يقيس للحكم الجمالي، إلا أنه في رأي أحد مؤلف هذا القياس يقيس (المساواة الجمالية) أيضاً والفرق بينه وبين اختبار جريفز فرق في المحتوى وليس في العمليات الأساسية المتضمنة.

وإذا كان اختبار جريفز وملري ينتميان إلى ميدان التفوق الفني فإن اختباري كنوبر وهورن ينتميان إلى ميدان الابتكار الفني. ويسمى أولهما Knauber Arc Ability ويصلح للمراهقين والراشدين، ويتطلب من المعهوس ما يأتي:

- 1- أن يرسم من الذاكرة أشكالاً في حدود المكان المسموح به.
- 2- أن يرسم شخصيات نمطية (بابا نويل مثلاً).
- 3- ترتيب مؤلف معين في حدود المكان المسموح به.
- 4- ابتكار وتكملة رسوم معينة من عناصر معطاة.
- 5- تحديد الأخطاء في أعمال مرسومة (مثل المنظور الخاطئ).
- 6- إنتاج أعمال فنية تدل على الخيال والحق والتمثيل الرمزي.

أما اختبار هورن فقد أعد خصيصاً للاستخدام في أغراض قبول طلاب كليات الفنون الجميلة، ويقيس على وجه الخصوص الابتكارية في الفن مستخدماً في كل سؤال وصفاً بسيطاً مرئياً أو مرشداً لكل رسم. ويتكلف الاختبار من قصيرين: الأول يتطلب من المعهوس أن يرسم 20 شيئاً مأخوفاً (مخل، كتاب... إلخ) في وقت قصير، كما يتطلب منه إنتاج رسوم تجريدية بسيطة مبتدئاً بمجموعة من المثلثات والمربعات وغيرها. أما القسم الثاني فيتطلب مستوى أعلى من الابتكارية وفيه يحل المعهوس 20 بطاقة مستطيلة على كل منها



خطوط تعيد كيدلية لمصورة يفتجها المفحوص. ويحكم على الرسوم على أساس  
الخيال الابتكاري من ناحية والخصائص الفنية من ناحية أخرى.

ومن الاختبارات التي تجرى لقياس القدرات في الفنون البصرية اختبارات تكميل  
الأشكال وتكميل الصور والتقدير الجمالي التي أعدها الدكتور محمد إسماعيل  
لقياس عاملي الطلاقة والتقدير الجمالي. ويقاس اختبار تكميل الأشكال عامل  
الطلاقة وفيه يطلب من المفحوص أن يضيف إلى خطوط مرسومة بضعة  
خطوط أخرى بحيث تجعل منه صورة لشيء ماكوف. ويقاس اختبار تكميل  
الصور عامل الطلاقة أيضاً وفيه يطلب من المفحوص أن يكتب في مكان معين  
من صورة عامل مرسومة أكبر عدد من الأشياء التي يمكن رسمها. أما اختبار  
التقدير الجمالي فهو يتتبع مصري لاختبار ماير، الذي أشرنا إليه. وقد حسب  
صدق الاختبارات بإيجاد معاملات الارتباط بين كل منها وبين المحك (وهو  
تقدير المدرسين للقدرة الفنية) فكانت النتائج كالآتي: 0.44 لاختبار تكميل  
الأشكال، 0.34 لاختبار تكميل الصور، 0.44 لاختبار التقدير الجمالي (ماير).

#### الختبارات لتفكير الابتكاري

من أهم خصائص التقويم النفسي المعاصر زيادة الاهتمام بقياس  
الابتكارية مسايرة للنشاط المتزايد في مختلف ميادين علم النفس لدراسة طبيعة  
النشاط الابتكاري والعوامل المؤثرة فيه. ولا يشع للمقام للإضافة في هذه  
المسائل، ويمكن الرجوع إلى بعض المصادر المتاحة في اللغة العربية ونستطيع  
قول بإيجاز أن الاهتمام بالتفكير الابتكاري يعبر عن حاجة المجتمعات  
المعاصرة لزيادة وتنمية ثروتها البشرية من الطعام والمهندسين والقادة الإداريين  
والقانونيين. ويرى كثير من العلماء للهنئين بهذا الميدان أن الابتكار حلوة

أساسية في هذه الأحوال جميعاً، فهو كما يقول تاولر (يختلف في العمق والمجال وليس في النوع) ومن رآه عدم التمييز بين الابتكار العلمي والابتكار الفني لأن الابتكار (يتضمن نظرة للمشكلات أكثر أساسية من التدريب المهني).

ويعود الفضل إلى المعلم الأمريكي المعاصر جيلفورد في زيادة حركة قياس التفكير الابتكاري، فقد أثرت نظريته عن تنظيم العقل في معظم بحوث الابتكار، وفيها يقترح تصنيف القدرات العقلية إلى فئتين كبيرتين هي قدرات التفكير، وقدرات الذاكرة. وتنقسم عوامل التفكير إلى ثلاثة أنواع هي:

1- قدرات التفكير المعرفي (بسطى الاكتشاف).

2- وقدرات التفكير الإنتاجي.

3- قدرات التفكير التقويمي.

أ- قدرات التفكير الإنتاجي التقاربي.

ب- قدرات التفكير الإنتاجي التباعدي.

والإنتاج التقاربي (أو الاختزالي) يتضمن تسبيق الاحتمالات عند إنتاج

إجابة واحدة محتملة للمشكلة، مثل: طويل إلى قصير مثل مرتفع إلى ... ؟ وبالطبع فإن أغلب اختبارات الذكاء العام تقيس هذا النمط من التفكير. بل إلى معظم الاختبارات العقلية التي أشرنا إليها في الفصول السابقة من هذا الكتاب تنتمي إلى فئة الإنتاج التقاربي. أما التفكير التباعدي فينطبل إنتاج أكبر عدد ممكن من الإجابات. مثل: ما هي الاستخدامات التي يمكن أن تفكر فيها لقلب للطرز ؟.

ويرى جيلفورد أن الإنتاج القواعدي صعب هام في التفكير الابتكاري. ولستعمل ووضع اختبارات جديدة لقياس هذه القدرة في مستويات مختلفة وفي محتويات معينة. ومن رأيه أن بعض عوامل الفكرة والتفكير المعرفي والتفكير التقويمي والتفكير الإنتاجي التقاربى ترتبط بالعمل الابتكاري.

ويستند على وجه الخصوص أن الموعية الابتكارية يمكن قياسها بالاختبارات التي تقيس قدرات الأصالة والطلاقة والمرونة من مختلف الأنواع، وقدرات إعادة التحديد والتضمين والتفكير التقويمي أيضاً.

وقد أثرت نظرية جيلفورد في ظهور بطاريتين رئيسيتين لقياس التفكير الابتكاري إحداها أعداها جيلفورد وزملاؤه بجامعة كاليفورنيا. والأخرى أعداها نورمن بجامعة مينيسوتا.

#### الختبارات لجيلفورد للتفكير الابتكاري

تتألف بطارية اختبارات التفكير الابتكاري لجيلفورد على النحو الآتي:

1- اختبارات جيلفورد للطلاقة وتتألف من 4 اختبارات يطلب فيها من

المفحوص أن يكتب كلمات بأسرع ما يمكن تحت الشروط الآتية:

أ- طلاقة للكلمات (الإنتاج القواعدي لوحدات الرموز)، يطلب فيه من المفحوص أن يكتب كلمات ذات خصائص معينة، كأن تتضمن حرف (ب) مثلاً.

ب- طلاقة الأفكار (الإنتاج القواعدي لوحدات المعاني): يطلب فيه من

المفحوص أن يذكر أسماء الأشياء التي تنتمي إلى فئة معينة، مثل  
سراويل التي تحترق.

ج- طلاقة للتداعي (الإنتاج التبادلي للعلاقات بين المعاني): يطلب فيه من المفحوص إعطاء كلمات تتشابه في المعنى مع كلمة أخرى.

د- الطلاقة التعبيرية (الإنتاج التبادلي لمنظومات لرموز): يتطلب فيه من المفحوص تكملة جمل، كل منها يتألف من 4 كلمات، بحيث تبدأ كل كلمة بحرف معين.

2- اختيار الاستخدامات البديلة (الإنتاج التبادلي لفئات المعاني): يطلب فيه من المفحوص أن يذكر الاستخدامات المحتملة لشيء معين غير استخدامه المألوف (قلب للظوب مثلاً) يستعمل في البناء استخداماً مألوفاً.

3- اختبار المترقيات (الإنتاج التبادلي لوحدة المعاني، والإنتاج للتبادلي لتحويلات المعاني) وفيه يطلب من المفحوص أن يذكر للتتابع المستقلة المترتبة عن حدث افتراضي (مثل: ماذا يحدث لو أن الناس توقفوا عن النوم؟). ويعطى للمفحوص في هذا الاختبار درجتان: إحداهما للاستجابات الواضحة وتيسر طلاقة الأفكار أو الإنتاج التبادلي لوحدة المعاني، والأخرى للاستجابات البعيدة وتيسر الأصالة، والإنتاج التبادلي لتحويلات المعاني.

4- الأعمال المحتملة (الإنتاج التبادلي لتضمينات المعاني)، ويطلب فيه من المفحوص أن يذكر الأعمال المحتملة التي يمكن أن يرمز لها بشعار (مصباح كهربائي مثلاً).

5- عمل الأشياء (الإنتاج التبادلي لمنظومات الأشكال): ويطلب فيه من المفحوص أن يرسم أشياء محددة باستخدام مجموعة مطاة من الأشكال (كالبازلز أو المكعبات). وعلى المفحوص ألا يستخدم إلا الأشكال المطاة فقط.

- 6- الاستكشافات (الإنتاج للتباعدى لرحلات الأشكال): ويشبه اختبار عمل الأبناء إما عدا أن كل صفحة في الاختبار تشتمل على نفس الشكل (دوائر مثلاً فقط) ويطلب من المفحوص أن يرسم أكبر عدد من الاستكشافات بإضافة تفاصيل إلى كل شكل من الأشكال المعطاة.
- 7- مشكلات عيدان القباب (الإنتاج للتباعدى لتعويلات الأشكال): وفيه يطلب من المفحوص أن يستبدل عدداً معيناً من عيدان القباب (المرسومة على الورق في صورة خطوط مستقيمة) بحيث ينتج من ذلك عدداً من المربعات أو المثلثات.
- 8- اختبار الزخرفة (الإنتاج للتباعدى لتضمينات الأشكال): وفيه يطلب من المفحوص أن يزخرف صوراً لأبناء مألوفة بأكبر قدر من الرسوم المختلفة.

ويلاحظ أن الاختبارات السبع الأولى من هذه القائمة تتطلب استجابات لفظية، بينما تستخدم الاختبارات الأربع الأخيرة. محتوى من الأشكال أو الصور. وهذه الاختبارات الأحد عشر تمثل عينة صغيرة من مجموعات الاختبارات التي أنتجتها بحوث جيلفورد وزملائه وتلاميذه بجامعة جنوب كاليفورنيا لأكثر من عشرين عاماً. وقد أعدت معايير مبدئية لهذه الاختبارات في صورة متويات ومعايير جسيمة C- score (وهو معيار يلخص المستويات للنتيجة إلى 11 مستوى وهو في ذلك يشبه المعيار للتصاعبي ويختلف عنه في عدد المستويات). وفي معظم الأحوال اشتملت هذه المعايير لعينات من الراشدين أو طلاب من سن 15 + أو كليهما. وتصلح هذه الاختبارات بوجه عام لمستوى المرحلة الثانوية وما بعدها، رغم أن جيلفورد يذكر أن بعض اختباره تصلح لمستوى المرحلة الابتدائية (ابتداء من الصف الأول).

والمشكلة الجوهرية في اختبارات التفكير الابتكاري هي مشكلة التصحيح فهي تتطلب من المصحح تدريباً كافياً على اتباع تعليمات التصحيح ودراسة الأمثلة التي تتضمنها كراسات التعليمات. ورغم ذلك فإن عملية تصحيح هذه الاختبارات تستغرق وقتاً كبيراً.

وحسب جيلفورد ثبتت هذه الاختبارات بطريقة التجزئة النصفية. وتتراوح معاملات الثبات بين 0.60، 0.80 أما عن الصدق فقد اُخذ اعتماداً كبيراً على صدق التكوين الفرعي كما يتمثل في اتفاق الاختبارات مع نموذج العقل الذي يقترحه. ولا زالت هذه الاختبارات في حاجة إلى أنواع الصدق الأخرى وخاصة الصدق المرتبط بالمحتكات.

وقد قام أحد الدكاترة بإعداد بعض اختبارات جيلفورد باللغة العربية، فأعد اختبارين لقياس الطلاقة اللفظية (طلاقة الكلمات) هما اختبار الطلاقة اللفظية (1)، ويتطلب من المفحوص إنتاج كلمات تبدأ بحرف معين، واختبار الطلاقة اللفظية (2) ويتطلب من المفحوص أن ينتج كلمات تنتهي بحرف معين. كما أعد (3) اختبارات أخرى لقياس بعض عوامل التفكير الابتكاري الأخرى وهي: اختبار الطلاقة الفكرية (قياس عامل طلاقة الأفكار)، واختبار الاستمالات (قياس عامل المرونة العقلية) واختبار المترتبات (قياس عامل الأصالة). وقد حسب ثبتت هذه الاختبارات من أداء عينة عدها (120) تلميذاً بالمدراس الثانوية العامة بطريقة إعادة الاختبار لاختباري الطلاقة اللفظية فبلغ معامل الثبات 0.62، 0.66 وباستخدام لي حسب ثبتت اختبارات طلاقة التفكير والاستمالات والمترتبات طريقة التجزئة النصفية وبلغت معاملات الثبات 0.75، 0.69، 0.80 على التوالي.

وفي تقدير الصديق حسب الباحث معاملات ارتباط الاختبارات بالتصصيل المدرسي، وتراوحت هذه المعاملات بين 0.27 (الاستعمالات)، 0.58 (المتريبات)، ولم تتضمن كرامة بالتطبيقات بيانات عن معايير الاختبارات.

### الختبارات تورنس للتفكير الابتكاري

نشرنا إلى أن اختبارات جيلفورد هي محصلة بحوث علمية في مجال التفكير الإنساني وطبيعة الذكاء، أما اختبارات تورنس للتفكير الابتكاري التي نتناولها الآن فقد ظهرت لمواجهة الضرورات التربوية والعملية كجزء من برنامج طويل من البحوث تهتم بالختبرات التطويسية التي تساعد على تنمية الابتكار، وبعض اختبارات تورنس هي الواقع تعديل للأساليب التي استخدمها جيلفورد بالإضافة إلى أن الدرجات المنقطة من البطاريات الكاملة تعتمد على العوامل التي توصلت إليها بحوث جيلفورد وهي الملائمة والمرونة والأصالة وإعطاء التفاصيل. وفي هذا الصدد يذكر تورنس أنه لم يصب إلى وضع الاختبارات (نقية عاملياً) أو (بسيطة التكوين)، وإنما إلى التوصل إلى اختبارات تتضمن مواقف مماثلة لما تتطلبه العملية الابتكارية في الظروف الطبيعية المعقدة. ولذلك فإن كل اختبار قد يصحح في ضوء عاملين، أو ثلاث عوامل، أو العوامل الأربع جميعاً.

وتتألف اختبارات تورنس للتفكير الابتكاري من (10) اختبارات، مصنفة إلى بطريقتين: بطارية لفظية وبطارية مصورة. وتسمى البطارية الأولى التفكير الابتكاري بالكلمات، والثانية: التفكير الابتكاري بالصور. وحتى يمكن التخفيف من التهديد الذي تتضمنه كلمة (الختبار) استخدمت في كل الأحوال كلمة (نشاط)، كما تؤكد التطبيقات على (اللاعب)، وتوصف هذه الاختبارات بأنها

تصلح لمختلف المستويات ابتداء من مرحلة الحضالة حتى مستوى الدراسات العليا، بشرط أن تطبق لبدءاً وشغوراً على للمستويات الأدنى من الصف الرابع الابتدائي. وتوجد صورتان متكاملتان لكل بطارية من البطاريتين.

وفي التفكير الابتكاري بالكلمات: الصورة (أ) والصورة (ب)، تتألف الأنشطة الثلاث الأولى (أسل وخن) من صورة غير عادية يستجيب لها المفوض بكتابة:

- أ- جميع الأسئلة التي يحتاج إلى توجيهها لكي يعرف ما يحدث في الصورة.
- ب- الأسباب المحتملة لما يحدث في الصورة.
- ج- النتائج المترتبة على ما حدث في الصورة.

أما النشاط الرابع فينتقل بطرق تحسين لعبة من لعب الأطفال بحيث تصبح مصدراً لمزيد من السرور والفرح لمن يلعب بها، ويتعلق النشاط الخامس بذكر الاستخدمات غير العادية لشيء ما، (علب الكرتون) في الصورة (أ)، (علب الصفيح) في الصورة (ب). ويتعلق النشاط السادس بالأسئلة غير المعتادة التي يمكن توجيهها عن الشيء المتضمن في النشاط الخامس. أما النشاط السابع فقد أهد على نسق اختبار المترقيات عند جيلفورد، ويطلب من المفوض أن يذكر كل ما يمكن أن يحدث إذا نشأ موقف غير ممكن الحوث. وتطلى البطارية درجات كلية في العوامل الثلاث: للطلاقة، المرونة والأسئلة.

ويتألف اختبار (التفكير الابتكاري: الصورة أ، والصورة ب من ثلاث أنشطة:



1- تكوين الصورة وفيه يعطى للمفحوص ورقة ملونة ذات شكل ملطي وعليه أن يضمها على صفحة بيضاء في الموضع الذي يشاء ويستخدمها كهداية لرسم صورة غير عادية (أو الورقة الملونة المستخدمة في الصورة أ) زرقاء للون، والمستخدم في الصورة (ب) صفراء للون في الطبعة العربية).

2- تكملة للصورة: ويتألف من (10) خطوط مختلفة يستخدم كلاً منها في رسم صورة منفصلة، وهذا الاختبار يستخدم في نفس الأسلوب المستخدم في اختبار هورن للاستعداد الفني الذي أشرنا إليه، مع فرق هام هو أن تعليمات اختبار تورس تؤكد إنتاج الأفكار غير المعتادة، كما أن التصحيح يعتمد على جوانب الابتكارية وليس الخصائص الجمالية والفنية.

3- يعطى النشاط الثالث للمفحوص قائمة من الخطوط المتوازية (الصورة أ) أو الدوائر (الصورة ب)، وعليه أن ينتج أكبر عدد من الصور والرسوم.

ويحصل للمفحوص في البطارية للصورة على (4) درجات هي:  
الطلاقة السرونة والأسئلة وإعداد التفاصيل.

وقد ظهرت هذه الاختبارات في عام 1971 على النحو الآتي:

- 1- التفكير الابتكاري بالكلمات: الصورة أ.
- 2- التفكير الابتكاري بالكلمات: الصورة ب.
- 3- التفكير الابتكاري بالصورة: أ.
- 4- التفكير الابتكاري بالصورة: ب.

- 5- كراسة تعليمات اختبارات تورنس للتفكير الابتكاري.
- 6- استمارة تصحيح (1) لاختبارات التفكير الابتكاري بالكلمات (الصورة أ، والصورة ب).
- 7- استمارة تصحيح (2) لاختبارات التفكير الابتكاري بالصور (الصورة أ، والصورة ب).
- 8- استمارة تقدير المعلمين للابتكار.
- 9- استمارة تقدير التلاميذ للابتكار.

وقد طبقت هذه الاختبارات على عينة كبيرة من تلاميذ المرحلتين الإعدادية والثانوية في العام الدراسي (1972 - 1973)، ويقوم الباحثان في الوقت الحاضر بإعداد معايير التصحيح، وتحديد المعالم السيكومترية تمهيداً لنشرها في كراسة منفصلة تمهيداً لنشرها، وقد صدرت الكراسة الأولى عام 1973 وتتضمن نظير الصورة أ من اختبارات الأشكال أو الصور، طبق الاختبار على 313 تلميذاً وتلميذة من مدرستين إعداديتين بالقاهرة تتراوح أعمارهم بين 12، 16 سنة، ولحساب الصدق استخدمت استمارة تقدير المعلمين للابتكار حيث طلب من مدرس الفرق التي أجريت عليها الاختبارات أن يختار كل منهم (5) تلاميذ يقعون في الفئة العليا، (5) تلاميذ يقعون في الفئة الدنيا من كل بعد من أبعاد الابتكار المقيسة: الطلاقة، المرونة، الأصالة، التفصيل. واختير كل تلميذ يتفق عليه مدرسان على الأقل. وحسبت دلالة الفروق بين متوسطات مجموعات التلاميذ الأعلى والأدنى في مختلف أبعاد الاختبار باستخدام مقاييس (ت) فوجدت الفروق في معظم الأحوال (باستثناء التفصيل) دالة وفي الاتجاه الصحيح إلا أن المستوى الدلالة كان عالياً (0.01) بالنسبة إلى بعد المرونة، وكان منخفضاً نسبياً (0.10) بالنسبة إلى بعد الطلاقة والأصالة.

ويستخدم في حساب الثبات الطرق الآتية:

- 1- حساب معاملات الارتباط بين متغيرات الاختبار (الطلاقة، المرونة، الأصالة، التفصيل، الدرجة الكلية) فوجدت المعاملات في كل الأحوال ذات دلالة إحصائية عند مستوى (0.01).
- 2- حسب معاملات الارتباط بين مصححين قلماً بتصحيح (38) اختباراً كل على حدة، وكانت معاملات ثبات التصحيح (0.99) للطلاقة (0.99) للمرونة وللأصالة، (0.97) للتفاصيل، (0.99) للدرجة الكلية ويرجع ارتفاع المعاملات على هذا النحو إلى التزام المصححين بقواعد مقننة في التصحيح، كما لجأ الباحثان في تدريب المصححين إلى طريقة تورنس، وتلكت بالفعل جدراً.
- 3- حساب الثبات بطريقة إعادة الاختبار وبلغت المعاملات 0.54 للطلاقة، 0.36 للمرونة، 0.53 للأصالة، 0.40 للتفاصيل، 0.50 للدرجة الكلية في الاختبار.

وقد أعد الباحثان معايير الاختبار في صورة درجات ثنائية، كما أعد للنماذج التجريبية لتصحيح الصورة (أ) من اختبارات الأشكال، وهي نماذج مشتقة من استجابات عينة للفنيين المصرية ذاتها وليست منقولة عن كراسات التطويرات الأجنبية.

# قياس الشخصية عن طريق الأحكام



## قياس الشخصية

قياس الشخصية عن طريق الأحكام والملاحظة المنظمة:

رأينا في قياس الشخصية عن طريق التقدير الذاتي أن الشخص يختار ويحطي تقديرات يرى أنها تعبر عن نفسه، أما عند قياس الشخصية عن طريق الأحكام والملاحظات المنظمة فإن الذي يقوم بتقدير السلوك أو الأداء المميز للفرد أو للمفحوص شخص أو لشخص آخر، ويمكن أن نقسم هذا النوع من قياس الشخصية إلى قسمين:

أ- مقاييس التقدير وتقديرات لقياس الاجتماعي.

ب-ملاحظة المنظمة لميزات السلوك.

وسوف نتناول كلا من هذين القسمين بشيء من التفصيل:

أ- مقاييس التقدير وتقديرات لقياس الاجتماعي:

يمكن أن نصف مقاييس التقدير وتقديرات القياس الاجتماعي إلى نوعين على أساس من يقومون بهما، النوع الأول يقوم به الرؤساء أو الكبار، بينما يقوم الأفراد أو الزملاء بالتقدير ولحكم في النوع الثاني.

1- تقديرات الرؤساء أو الكبار:

نحن نلاحظ أن الرؤساء أو المدرسين مثلاً يقومون باستمرار بتقديم نوع من الأوصاف لمرووسهم أو لمن هم في كفالتهم، إلا أننا لا نستطيع أن

من هذه الأوصاف لأنها تختلف في أسلوبها، لهذا استعملت مقاييس الأوصاف  
في يتوفر قدر مقبول من التشابه بين التقديرات.

ويتكون مقياس التقدير من قائمة من السمات يجري عليها التقدير.  
ومن المقاييس من قائمة بسيطة من الصفات يتم الاختيار بينها، أو قد و  
بأساً متصلاً يتضمن كلاً وصفيّة متحدة للاختيار منها. ويوضح الآ  
لى، هذا النوع من مقاييس التقدير المتصلة:

<p>وقائع مؤيدة</p>	<ul style="list-style-type: none"> <li>• يسمى إليه الآخرون.</li> <li>• يحبه الآخرون.</li> <li>• يحمله الآخرون.</li> <li>• يتجنبه الآخرون.</li> <li>• لا مجال لملاحظته.</li> <li>• يسمى إلى إضافة أعمال جديدة إلى نفسه.</li> <li>• يكمل للعمل الإضافي المقترح.</li> </ul>	<p>1- كيف تتأثر أنت والآخرون بظهوره وسلوكة ؟</p>
<p>وقائع مؤيدة</p>	<ul style="list-style-type: none"> <li>• يعمل واجباته العادية تلقائياً.</li> <li>• يحتاج إلى تحريك أحياناً.</li> <li>• يحتاج إلى تحريك كثير ليعمل واجباته العادية.</li> <li>• لا مجال لملاحظته.</li> </ul>	<p>2- هل يحتاج إلى تحريك أو أنه يعمل من تلقاء نفسه ؟</p>

نلاحظ أن مقاييس التقدير عن طريق الرؤساء تكمن فيها مصادر متعددة للخطأ منها مثلاً ما يسميه كرونباك (خطأ الكرم) ويقصد به ميل من يقوم بالتقدير إلى إعطاء تقديرات طيبة موبدة، فإذا مثل مدرس مثلاً في قدر تلميذاً ما فيما يختص بالنيل إلى التعاون، فإنه سوف يقدّر كل التلاميذ الذين لا يشيرون للشغب تقديراً عالياً جداً، في كسبي طرف المقاييس، بينما يهبط بتقدير لثقة من التلاميذ مثوري للشغب، ويكون الارتفاع في تقدير الغالبية رد فعل تمييزهم عن مثوري الشغب. ومن أسباب خطأ الكرم أيضاً أن الرئيس المقدر قد يحس أن مما يمس قدرته الرئاسية أن يعترف أن مرؤوسيه ليسوا أكفاء، أو أن الرئيس قد يجد من الأسر عليه أن يحكم أحكاماً طيبة من مرؤوسيه إجمالاً يدل أن يتولى ويطلق النظر ليميز بينهم تمييزاً دقيقاً.

ومن مصادر الخطأ أيضاً في هذا النوع من التقدير للمفوض والإيهام، فقد يفسر المقدر سؤالا أو مصطلحاً تفسيراً مختلفاً، فقد يفسر شخص (القيادة)، مثلاً على أنها تركيز السلطة في يد القائد، والافراد بالتخاذل لقرارات، فعندما يقدّر هذا الشخص المفوضاً تتوالى فيه هذه الصفات فله يقدّره تقديراً عالياً ومخالفاً لتقدير شخص آخر للمفوض ذاته، لأن للمقدر الأخير فهم القيادة على أنها تشجيع للمرؤوسين ومساعدة على اتخاذ قرارات جماعية.

ويشتهر أسلوب الاستجابة عند المقدر مصدراً آخر من مصادر الخطأ في مقاييس التقدير بواسطة الرؤساء أو الكبار، فعلاً نجد أن مقدار ما يندرج في متصل أطراف المقاييس في وصف المفوضين، بينما يميل شخص آخر إلى الإسراف في استعمال هذه الأطراف، أي يميل إلى تصنيف المفوضين إلى



لقسام متميزة جداً. ومن ثم نجد أن مثل هذين المقدارين يختلفان عندما يقدران شخصاً واحداً.

ومما يؤثر في تقديرات الرؤساء أيضاً أن مطوماتهم عن المفحوصين قد تكون محددة، فقد يختلف تقدير مدرس للتربية الفنية ومدرس اللغة العربية في تقدير رنود فعل ثمبذ معين للهوطه ويرجع الاختلاف إلى أن كلا منهما يعرف التلمبذ وبلاظه في عدد محدود من المؤلف، وعلى هذا تكون معرفته به محدودة ولهمت شاملة.

ومن مصادر الخطأ في هذا النوع من التقدير (نثر الهالة) أو (وقع الحلة). ويحدث هذا الأثر أو للوقع عندما يتأثر المقدر تأثيراً قوياً بالانطباع العام عن المفحوص، فيتكون عنده رأي عام أو تقدير عام عن قيمة الشخص يؤثر كثيراً في تقديره لسمات هذا الشخص الخاصة. وهذا النوع من الخطأ يجعل نمط السمات في لف رد هائناً حد المقدر.

إلا أنه من الممكن تلافي كثير من مصادر الخطأ هذه إذا راعينا النقاط التالية:

- 1- الاختيار السليم لمن يعين بالتقدير، من حيث قدرته ودرجة انتماجهم وتأثرهم شخصياً بتقديرهم للمفحوصين، فقد سبقت الإشارة إلى أن المقدر قد يميل في تقديره إلى التحيز لصالح رؤوسه لأن في هذا تأكيد لكفاءة رئاسته.
- 2- كلما زادت معرفة المقدر بالمفحوصين كانت تقديرته أكثر صديقاً.

3- إذا استسلنا مقراً بوجه أكثر من مرة فمن المفيد أن نحتفظ بتقديراته لمقارنتها وللتعرف على أسلوب استجابته، أو ما يمكن أن نسميه خطأه الثابت.

4- من الممكن زيادة ثبات مقاييس التقدير عن طريق الاعتماد على أكثر من مقدر واحد، على أن يكونوا على معرفة كافية بالمفحوصين.

5- مما يزيد من قيمة مقاييس التقدير أن تعد بطاوة ودقة، وبصفة عامة نجد أن المقولس ذا الخصص نقط يعطي تمييزاً أكثر دقة من المقولس أو التلقمة التي تعتمد على إجراءات لعم أولاً. كما أن مقولس الخمس نقط يذب إلى أنواع متبيلة من الاختلافات والأكراهات. وكذلك يساعد الاختيار الإجابري الذي نلكننا حه في بعض اختبارات التقرير لذاتي، على زيادة قيمة مقولس التقدير. والهدف من الاختيار الإجابري هو فصل عمل المقدر فيما يتصل بوصف ما بعمله المفحوص عن تقرير هذا للعمل، أي أن المقدر يكون مسؤولاً عن وصف المفحوص، بينما يقوم المسئول عن اتخاذ القرار بتقرير هذا الوصف. وهذا النوع من مقاييس التقدير التلقم على الاختيار الإجابري أكثر ملائمة لقرارات المؤسسات فيما يتصل بالانتقاء والتصنيف، ولكنها لا تناسب كثيراً صلية التوجيه التربوي أو الوصف التفصيلي للفرد.

لما عن صدق مقاييس تقدير الرؤساء فتشير الدراسات التي أجريت عنها إلى ضرورة أخذ هذه المقاييس بحذر فيما يتصل بالحكم على الأفراد من جهة، وبإستعمالها للتحقق من صدق بعض مقاييس واختبارات الشخصية من جهة أخرى. ومن أمثلة هذا أن طريقة تقدير الكبار لمتعلمات في دراسة صدق مقولس المسئولة الاجتماعية، للصورة (ت)، الذي تناوله الفصل السابق، حيث

قام المدرسون بتقدير تلاميذ العيلة على مقياس خالص بالمدرسين وكان معامل الارتباط بين تقدير المدرسين ودرجات التلاميذ على المقياس 0.46 وقد اعتبر هذا المعامل مقبولاً (وبخاصة إذا أخذنا في الاعتبار أن تقدير المدرس للتلميذ في أي ناحية من نواحي السلوك تتدخل فيه عوامل كثيرة، منها على سبيل المثال، تحصيل التلميذ الدراسي، ومستواه الاقتصادي الاجتماعي، ودرجة جلالية التلميذ الشخصية، وتفضيل المدرس وتحيزه الشخصي وهكذا).

وقد أثبتت دراسة قام بها بيترز وآخرون أن ارتباط تقديرات الرؤساء لمدى معرفة العاملين درايتهم بأعمالهم مع هذه المعرفة والدراية كما قيست باختبار خالص، كان الارتباط بينهما 0.35 بينما وصل هذا الارتباط في بعض الأقسام 0.55، كما أظهرت دراسات أخرى ارتباط التقديرات بطول المدة التي يعرف فيها المقرر المفحوصين، ويمتاز حبه لهم.

ولكن مقياس تقدير الرؤساء والكبار، على الرغم من حدودها وتعتبر مصدرأ غنياً بالمعلومات التي لا يمكن الحصول عليها إلا من طريقها.

## 2- تقدير الأقران:

هي التي تعطي معلومات تطور في قسمها على تقديرات الكبار أو الرؤساء، وحتى عندما تتوفر تقديرات الرؤساء فإن تقديرات الأقران والزلاء تعطي بيانات عن جالب مختلف من الشخصية. فإذا كان رئيس واحد هو الذي يقوم بالتقدير فإن قيام مجموعة من الأفراد أو الزلاء بالتقدير سوف يعطي فرصة لعدد كبير من الأحكام مما يترتب عليه أن متوسط التقديرات بالنسبة إلى سمة معينة من سمات الشخصية سوف تكون أكثر ثباتاً من التقدير الذي يقوم به

رئيس واحد. ومن الممكن أن ننظر إلى تقديرات الأفراد على أنها تعبير موضوعي عن سعة الفرد، فالمسعة تقود إلى حد ما على السلوك الفعلي، إلا أن نمط العلاقات الاجتماعية في داخل الجماعة لا شك يؤدي إلى وجود نوع من التحيز.

ومن الأساليب المستعملة في تقدير الأفراد ما يسمى أسلوب تعيين الأسماء أو الترشيح، ويستعمل هذا الأسلوب كثيراً في القبلات الاجتماعي وفي هذا الأسلوب يسأل كل عضو من أعضاء الجماعة أن يسمي عدداً محدداً من الأشخاص الذين يتميزون بشكل واضح في ناحية من النواحي، مثل القيادة، كما يسمى أيضاً الأشخاص الذين يتقصبهم هذه الجماعة. وبالنسبة إلى الأطفال فإن صلية للتسمية قد تتخذ صورة لعبة التخمين، أو اختبار (خمن من) الذي يصف الأدوار المختلفة التي يمكن أن يلعبها الأطفال، ويسمى كل عضو في الجماعة أسماء الأطفال الذين يعتقدون أن كل وصف يلائمهم. وترسم لكل طفل صفحة نصية عن طريق عدد المرات التي تشير إليه فيها لكل وصف.

وإذا كان من الممكن دراسة خصائص الفرد بطريقة خمن من، فلي الرسم الاجتماعي (الموسيوجرلم) يعطى فهماً واستبصاراً أعمق في تمييز التجمعات الداخلية في الجماعة، وبين البناء الهرمي للقيادة والمنزلين، أي هذا الرسم الاجتماعي يعتبر طريقة لدراسة البنية الاجتماعية في الجماعة.

والذي أنشأ الرسم الاجتماعي وطوره ماريلو، إلا أن الأسلوب قد عدل بعدة طرق قللت من فاعليته في سبيل السهولة والبصر في الاستعمال. والفضل طريقة في استعماله هي أن يسأل أعضاء الجماعة أن يختاروا رفقاء لهم في نشاط معين. ولا بد أن تكون التطلعات متصلة بالنشطة حقيقة للجماعة، كما يجب

أن تكون الاختبارات حقيقية ويجب أن يعلم أعضاء الجماعة أن المعلومات التي يطونها سوف تصان سرّيتها، وأن البيانات للمعطاء سوف تستخدم فيما أخطر أعضاء الجماعة أنها سوف تستخدم فيه، مثلاً لاختيار لجان، أو جماعات عمل فرعية، أو تنظيم الجلوس في فصل دراسي وهكذا.

ويستند لرسم الاجتماعي الذي نحصل عليه على السؤال الذي نوجهه، فإن نمط الرسم الذي ينتج عن سؤال مجموعة من التلاميذ عن لغيراتهم لأقرانهم الذين يحبون أن يستذكروا معهم سوف يختلف عن نمط الرسم الناتج عن سؤال عن اختبارات من يحبون لسطحاهم في رحلة.

ويشمل مصطلح التقدير بالقياس الاجتماعي كل طرق تمييز العلاقات الاجتماعية بين أعضاء الجماعة. وليس هناك فصل قاطع بين تقدير الأفراد الوصفي، وبين التقدير بالقياس الاجتماعي، إلا أننا نجد أن النوع الأخير بصفة محددة بأسئلة صحت بحبهم المقدر أكثر أو يفضل للعمل معهم، ولهذا فهي تتصل باستجابة المقدر كما تتصل بشخصيته.

ولتقديرات الأقران لستعمالات متعددة منها أن فائد للجماعة يستعملها في التعرف على الأشخاص الذين يحتاجون إلى عناية خاصة، أو يستعملها في إعادة تنظيم الجماعة حتى تؤدي وظيفتها بشكل أفضل. كما أن هذه التقديرات يمكن أن تتخذ أساساً للانتقاء والتصنيف، وهناك من يعتبرها لنقي مقياس للقيادة، لما بالنسبة إلى المرشد النفسي، فإن تقديرات الأقران توضح خصائص للشخص التي تعرق قبول الجماعة له، وخاصة عندما ينقص الشخص الاستبصار الكافي في رأي الآخرين فيه.

ويذهب كثير من الباحثين إلى أن تقديرات الأقران أثبتت أنها إحدى وسائل التقدير التي يمكن أن يوثق بها إلى درجة بعيدة لعدة أسباب هي:

1- أن عدد المقدرين كبير، وقد أشرنا منذ قليل إلى أهمية هذه في زيادة ثبات التقديرات بصفة عامة.

2- أن أقران الشخص عادة ما يكونون أكثر قدرة من غيرهم على ملاحظة السلوك المميز له، وهم من ثم أفضل في الحكم على سمات معينة تتصل بالعلاقات الاجتماعية من المدرسين أو المشرفين أو أي ملاحظين خارجين.

3- أن آراء أعضاء الجماعة الآخرين، سولياً كانت أم خطأ، تؤثر في سلوكهم وتصرفاتهم وبالتالي فإن هذه الآراء كتحد جزئياً طبيعة تفاعلات الشخص مع الجماعة. فمثلاً، إذا كان أعضاء جماعة يرون أن أحد أعضاء جماعتهم، ويمكن اسمع طلياً مثلاً، رجل يمكن الاعتماد عليه والثقة به، فإن هذا الرأي يوجه ويحدد نوع تعاملهم مع علي، وغالباً ما يحدد هذا العامل نوع السلوك من جهة علي كما يحدد نوع التفاعل من جهة الجماعة. فالجماعة تعامله على أنه ثقة، وهو بذلك تجاه الجماعة وفق توقعاتها منه ولا شك أن هذا يعطي تقديرات الأقران صفة أقرب إلى الواقعية.

من الممكن أن نسلينا للملاحظات المنظمة وصفاً أكثر دقة للسلوك أو الأداء المميز إذا قورنت بأساليب التقرير الذاتي أو الأحكام وتقديرات التي يصدرها الآخرون عن الشخص، ذلك لأنه في التقرير الذاتي من المحتمل أن يتدخل للتذكر الانتقائي، بل أنه يحدث أيضاً عندما نستعمل أساليب التقدير

والحكم المختلفة، وكذلك نجد أن الملاحظة في هذه الأساليب الأخيرة ليست قائمة على الاختيار المنظم للمواقف التي يلاحظ فيها السلوك، بل إنها تقوم على ملاحظة المواقف كيفما اتفق وقوعها، وفي هذه الأساليب أيضاً لا يلاحظ الشخص في كل المواقف. لهذا قلنا إن الملاحظة المنظمة تعطي وصفاً أدق للسلوك المميز.

ومن السهل أن نلاحظ سلوك مفحوص ما في ظروفه العادية، وهذا ما يسمى بالملاحظة المبدئية، هذا النوع من الملاحظة يكون أكثر ملائمة في بعض الدراسات من الملاحظات المقتنة، فكثير من الباحثين يرى أنه من المستحيل أن نفهم الشخصية إلا إذا لاحظنا المنحصر وهو يستجيب إلى الظروف التي لها مغزى بالنسبة إليه، لأن المثيرات المختلفة لها معانٍ ومغاير مختلفة بالنسبة إلى الأفراد المختلفين. وعلى هذا فهم يرون أن المواقف المقتنة، حيث تلاحظ ردود أفعال الأشخاص لموقف بعينه، تبرز أنماط السلوك الهامة كما تبرزها الظروف العادية، غير المتشابهة، التي يعيش فيها الفرد. إلا أن المشكلة هنا تكمن في كيف نرى قدراً كافياً من سلوك الفرد، والحصول على تسجيلات لملاحظات السلوك يمكن الاعتماد عليها. أي أن المشكلة هنا ذات شقين: الأول، عينة السلوك وما يرتبط بها من صعوبات، والثاني، عيوب أو أخطاء من يقوم بالتسجيل، ولنتكلم عن كل منهما بشيء من التفصيل.

#### لختيار عينة السلوك

لكي نعرف السلوك أو الأداء المميز لشخص ما لا بد أن نعرف ما يميز تصرفه من موقف بعينه، ولكن للمواقف تتغير من يوم إلى يوم ومن لحظة إلى أخرى، والطريقة الوحيدة للعثق إلى أن ما نلاحظه هو السلوك المميز هي أن

لدرس هذا الشخص في مواقف ومناسبات عديدة، وهو أمر مجهود ومكلف. فالدافع له لا بد أن نصل إلى نوع من التوفيق بين الاختيار الدقيق لجودة السلوك والاقتصاد في الإنفاق والجهد.

وأنه من التصير جداً أن نقوم باستنتاجات عن لروق بين الأفراد من ملاحظة سلوكهم في مواقف معينة، لأنه لا يمكن ملاحظة لرين في مواقف واحد بعينه، وحتى إذا كان للموقف واحد في ظاهره فإن الظروف السابقة الخاصة بكل منهما تجعلهما يسلكان سلوكاً مختلفاً، ولهذا فإن أقصى ما يمكن أن تقدمه لنا للملاحظات المقارنة هو بيان كيف يملك أشخاص مختلفون في ظروفهم الحالية، ولكن هذه الملاحظات لا تضمن لنا أن تستمر هذه الفروق عندما تتغير طبيعة هذه الظروف.

ومن أفضل الأساليب التي تؤدي إلى المقارنة الدقيقة ما يسمى بالعينات الموقوتة ولي هذا الأسلوب يخطط مسبقاً جدول أو قائمة للملاحظات ويصمم هذا الجدول بشكل عشوائي يضمن أن يلاحظ كل مفحوص في ظروف متشابهة مع الآخرين بحيث تمكن المقارنة. فمثلاً عند دراسة الاتصالات الاجتماعية للأطفال ما قبل المدرسة وضع جدول لملاحظات لمدة دقيقة واحدة، وبعد ملاحظة الطفل لمدة دقيقة كتب للملاحظة تقريراً كاملاً عما لاحظته من تفاعلات الطفل الاجتماعية. ولوحظ الأطفال وفق نظام مرسوم من قبل وكان يحل من يوم إلى يوم، وفي خلال الدراسة كان كل طفل يلاحظ عدد مرات متساوية في خلال الخمس دقائق الأولى من ساعات اللعب الحر، ثم الخمس دقائق الثانية وهكذا... والعينات الموقوتة قصيرة الأجل والموزعة توزيعاً حصناً عدة مزايا، فالصورة المتركمة من الملاحظات فيها تكون أكثر تشبهاً للسلوك من نفس القدر من



المعلومات التي يحصل عليها من ملاحظات من مرات لكل وذلك فترات طويلة، كما أن لخطأ الذاكرة تقل هنا لأن للملاحظ يقوم بالتسجيل أثناء وبعد للملاحظة مباشرة. والعينات الموقوتة أكثر مناسبة لتسجيل حقائق معينة يمكن أن تصاغ في صورة رقمية، مثل عدد مرات الاتصال الاجتماعي بالآخرين.

ومن الممكن الحصول على عينة أكثر اتساعاً بأسلوب (السجل اليومي) الذي أنشأه باركر وزملاؤه. كان هدفهم هو دراسة نمط حياة الطفل ومع الاهتمام بالخصائص بالمجالات التي يتحرك فيها.

والسجل اليومي مزياً أن يعطي صور عن تتابع للنشاط، ولكن له عيوبه، إذا كان للملاحظ معروفاً ومكتشوقاً تأثر السلوك الذي يلاحظه، والاعتماد على يوم كامل بعينه لا يكفي للحصول على معلومات مميزة كاملة عن المفحوص.

ومن المهم أن نلاحظ بالنسبة إلى عينة السلوك للملاحظة عن طريق لعينة الموقوتة ما يلي:

- 1- أن الاستجابات التي بعدها الملاحظ متشابهة أو واحدة قد يكون لها في الحقيقة معان مختلفة. وأن المواقف التي تبدو له متشابهة قد تستدعي استجابات متباينة.
- 2- أن الاستنتاجات المشتقة من مواقف واحد، حتى ولو كانت على أساس تراكم ملاحظات، تكون، أي الاستنتاجات مسمحة بالنسبة إلى هذا الموقف فقط.

3- أن الملاحظات المطلوبة والملائمة للحصول على معلومات ثابتة ويمكن الثقة بها، يعتمد هذا المبدأ على المشكلة، إلا أنه بصفة عامة نجد أن عدداً كبيراً من الملاحظات القصيرة أفضل من عدد قليل من الملاحظات التي تعتمد على فترات طويلة.

### عيوب من يقوم بالملاحظة

الواقع أن بعض عيوب الملاحظين قد يرجع إلى طبيعة المواقف الاجتماعية ذاتها التي تكون موضوع ملاحظة حيث تحول درجة تعقدها الكبير دون التقرير الشامل المستغرق لها. أما العيوب الأخرى فنرجع إلى الملاحظة نفسها.

وإذا كانت أخطاء الملاحظ مجرد حذف أو إغفال عشوائي لبعض ما يلاحظ فليس لهذا وزن كبير، أما ما يجب الاهتمام به فهو الأخطاء المنظمة، كان يبلغ باستمرار وبصفة مطرودة في تأكيد بعض أنماط الأحداث، كما ينفل الإكثار إلى بعضها الآخر.

عندما يلاحظ مراقبون موقفاً واحداً فإنهم يعطون تقديرات بدرجة كبيرة، ذلك لأن كل ملاحظ عده حساسية لأنماط معينة من السلوك أكثر من غيرها والوضع الأمثل هو أن يبني الملاحظ انطباعه على كل تصرف يمكن أن يكشف أو يعكس خصائص أسلمية، ولكن عندما يركز من يقوم بالملاحظة والمراقبة على شيء بعينه فإنه بالضرورة لا بد أن يخلل ملاحظة أشياء أخرى.

وثمة عيب آخر وهو أن الملاحظين يفسرون ما يرون، فإذا قام الملاحظون بتسجيل الحقائق الموضوعية، فقد يصل آخرون إلى تفسيرات مختلفة

شاماً، ولكن الناس يميلون إلى تفسير ما يرون، وعندما يفسرون فإنهم يميلون إلى تجاهل الحقائق التي لا تتفق مع تفسيرهم بل أنهم قد يخترعون حقائق لتكمل تفسيرهم.

ومما يقل من آثار هذه العيوب أن يتم التسجيل بشكل منظم، وأن يكون كلما أمكن، في صورة وحدات سلوكية يمكن عدّها فئلاً، من الممكن أن نصف مدى لفظنا صال مصنع في عملهم عن طريق التسجيل لأزمني للحظات التي يكونون فيها في عمل فطري، والحظات ينشغلون فيها عن العمل للحصول على الأذرات، أو زيارات زملائهم، أو يتعرضون فيها لمؤثرات للتشخيص. وحتى إذا كان السلوك فملاحظ على درجة كبيرة من اللتووع بحيث يصعب وضعه في صورة وحدات سلوكية فمن الممكن أن تحدد أقسام السلوك ونعد قبل ويكون عمل فملاحظ هو أن يصنف للسلوك الذي يلاحظه وفق هذا التقسيم، ويضع علامة في القسم المناسب لكل أسلوب يلاحظه كما هو الحال في أقسام بيلز.

وعلى الرغم من فائدة التسجيل والعد الموضوعي لوحدات السلوك وملائمته للبحوث الفصلية فإن فائدته محدودة في التوجيه والإرشاد النفسي.

أما السجلات القصصية أو الوصفية فإنها، وإن كانت تتفهمها دقة الموضوعية، فإنها تعطي وصفاً عاماً للشخص، لأن الملاحظ فيها يوجه اهتمامه إلى أي سلوك يبدو جدير بالملاحظة بدل أن يركز على سمات معينة للمفحوصين كالم إلا أن الملاحظ في السجل القصصي يصف بدقة ما يلاحظه فاصلاً بين الوقائع وتفسيراتها. وعندما تتراكم هذه السجلات القصصية فإنها تعطي صورة أكثر غنى من أي وسيلة أخرى على هذه الدرجة من البساطة. ومسؤولية الملاحظ الكبرى هنا تتركز في اختبار الوقائع التي تستحق للتقرير، وفي أن يتحرى الموضوعية بقدر الإمكان.

# **قياس الشخصية وسائل التقرير الذاتي**



## ومقال التفكير الذاتي

أشرنا فيما سبق إلى تقسيم المقاييس النفسية بصفة عامة إلى قسمين رئيسيين هما: مقاييس الأداء المميز، ومقاييس أقصى الأداء، وعرفنا أن في النوع الأول من هذه المقاييس يكون الهدف من قياس هو معرفة ما يقوم به للشخص عادة، بينما النوع الآخر من المقاييس يكون الهدف هو معرفة أقصى ما يمكن أن يقوم به الشخص في ظروف استثارة ودفعية غير عادية.

ويمكن أن نضع تحت القسم الأول، أي مقاييس الأداء المميز، اختبارات الشخصية، واختبارات السيول والاتجاهات بينما تدخل لاختبارات للقدرة بأنواعها المختلفة تحت القسم الثاني وهو اختبارات أقصى الأداء. وفي هذا الفصل نتكلم عن قياس الشخصية باعتباره قياساً للسلوك أو الأداء المميز. وسوف نتبع، بصفة عامة، التقسيم الذي حدده كرونباك لأنواع مقاييس الشخصية، فدرجها تحت هذه الأقسام الثلاثة:

- 1- قياس لشخصية عن طريق التقرير الذاتي، وهي ما سيتناوله الفصل الحادي.
- 2- قياس لشخصية عن طريق أحكام الآخرين. والملاحظات المنظمة، وهي ما سيتناوله الفصل الثامن.
- 3- قياس لشخصية عن طريق لاختبارات الأداء وهي ما سيتناوله الفصل العاشر.

## قياس الشخصية عن طريق التقرير الذاتي:

على الرغم من أن تعبير السلوك المميز الذي تكررت الإشارة إليه يحمل طابعاً (سلوكياً) فإن المذهب السلوكي لم يكن هو الذي أسهم في نشأة وتطوير مقاييس للشخصية وإنما يرجع الفضل في هذا إلى الاتجاه الظاهراتي (الفيثومونولوجي)، فهذا الاتجاه يتميز بأنه يركز على العالم كما يدركه الشخص، أو على العالم الذاتي للشخص، بما في ذلك إدراك الفرد لذاته، وعلى أن السلوك إنما يتحدد، ولا يمكن أن يفهم إلا في ضوء هذا العالم الذاتي للداخل، كما يذهب هذا الاتجاه إلى أن غالب مشكلات التوافق التي يتعرض لها الفرد ترجع إلى إدراكه للشخص للأحداث والأشخاص وليس لواقع هذه الأحداث أو الأشخاص في ذاتهم. وبالتالي فإن العلاج النفسي، وفق هذا الاتجاه يجب أن يركز على تصحيح إدراك الشخص.

شير إلى المدرسة الفيثومونولوجية في بداية كلامنا عن قياس الشخصية لأن أول مقاييس أو استفتاءات للشخصية كانت قد وضعت بهدف قياس هذا العالم الداخلي الذاتي من الإدراك والشعور. ففي العقد التاسع من القرن التاسع عشر وضع جالتون هذه المقاييس عندما احتاج إلى طريقة يمكن أن يطبقها على عدد كبير من مفوضيه أثناء دراساته للصور العقلية، ثم استعمل هذا النوع من الاستفتاءات بعد ذلك ستالي هول في دراساته للموسم عن نمو المراقمين ولكن الاستفتاء كانت له وظيفة خلسة بالنسبة إلى كل من جالتون وستالي هول، بالنسبة إلى جالتون كان التقرير الذاتي، يستعمل على أنه الوسيلة الوحيدة الممكنة للحصول على معلومات عن أمور ولحادث في الحقل المفحوص، في داخل رأسه، بينما كان التقرير الذاتي بالنسبة إلى ستالي هول يستعمل بهدف

تجنب الجهد الشاق والوقت الطويل الذي يمكن أن يستهلك في الملاحظة المباشرة لسلوك المراقبين.

لما أول استفتاء وضع بقصد قياس الشخص فقد كانت قائمة وردت البيانات الشخصية. وظهرت خلال الحرب العالمية الأولى عندما ظهرت الحاجة إلى تمييز أولئك الجنود الذين عندهم قابلية للتجهيز أثناء القتال، ولم يكن يتوافر عدد كاف من المتخصصين لتحقيق هذا الهدف. لهذا وضع وودورث قائمة من الأعراض التي يهتم الفاحص بالتعرف عليها لاستبعاد من تظهر عندهم، وعندما أعطيت القائمة منزلت بالفعل أولئك الذين قرروا وجود عدد كبير من الأعراض عندهم ثم خضعوا لاختبار أكثر دقة، وكانت لاختبار البيانات الشخصية هذا قيمة كبيرة في ذلك المجال.

ويعتبر اختبار وودورث الحد الأكبر لعدد كبير من استفتاءات للتوافق، وهي التي تتكون أساساً من قوائم من المشكلات أو الأعراض أو الشكاوى التي على المفحص أن يختارها. إلا أن مثل هذه القوائم لا تدعى أنها تعطي وصفاً دقيقاً عن الشخصية وهي غالباً ما تعطي درجة واحدة تمثل مستوى للتوافق. وهي تتكون من عناصر يمكن أن يميز بين أشخاص معروف أنهم غير متوافقين وبين آخرين معروف أنهم عاديون. والاستعمال الرئيس لمثل هذه المقاييس هو التعرف على أولئك الأشخاص الذين يحتاجون إلى رعاية وإرشاد نفسي، وعلى هذا يمكن أن ننظر إليها على أنها أدوات فرز تساعد في تحديد الأشخاص الذين يختارون أعراضاً وصفات وانفعالات تتصل بهم شخصياً.

ولا يمكن اعتبار استفتاءات للتوافق مقاييس دقيقة لسمات محددة، فالمعلومات التي تمسك بها مطومات ليس لها صفة العمق. لهذا فإنه في الفترة



بين 1920-1945، تفرعت النظرة إلى مثل هذه الاستفتاءات وذلك بتأثير المدرسة السلوكية واستحارها للتأمل الباطني الذاتي الذي هو أساس اختبارات التوافق، فوجب أن يكون المقياس بديلاً لملاحظة السلوك، وعلى هذا يجب أن تركز أسئلته بشكل أكبر على ما يفعله الشخص لا على ما يشعر به أو يفكر فيه. وعلى هذا اتسعت مجالات الاختبارات بحيث تصف أكبر عدد ممكن من جوانب الشخصية، وتلخص الاستجابات بحيث تصف أكبر عدد ممكن من جوانب الشخصية، وتلخص الاستجابات بإعطاء درجات لعدد من السمات أو الأبعاد الاجتماعية، ذلك لأنه كان ينظر إلى الشخصية في ذلك الوقت على أنها مجموعة أو حزمة من العادات، وكان يحكم على الشخص حسب قوة سمات مثل: الثقة بالنفس، الإصرار، سهولة تكوين صداقات... إلخ والسمات القوية هي السمات التي تصف استجابة يقوم بها الشخص عادة وبكثرة.

وقد كانت قائمة السمات أو أقسام السلوك التي تعطى لها درجات في الاختبارات تختار بشكل قسري ودون خطة معينة في الاختبارات الأولى، فمثلاً سمات مثل الثقة بالنفس كان مصدرها الخبرة اليومية، بينما تشتق سمات أخرى مثل الإكراه من نظرية من نظريات الشخصية. وقد ظهرت عشرات من الاختبارات بهذا الشكل، تأخذ اللاحقة منها من السابقة عليها بضمعة عبارات أو أسئلة وتضيف إليها وتصممها بحيث توضع في تجمعات سمات مختلفة. إلا أن أكثرها شيوعاً هو اختبار بيرنرويتز للأشخصية.

وتعطى فيه للدرجات للدلالة على النزعة العصابية، الاكتئاب الذاتي، الإكراه، والسيطرة. ويتكون الاختبار من (25) سؤالاً يجيب عليها المفحوص بعدم أو لا. وقد وضعه بيرنرويتز على أساس أن سلوك الفرد في مواقف واحد

قد يكشف عن سمات شخصية مختلفة، وعلى هذا يمكن أن ينظر إلى كل سؤال من أسئلة الاختبار على أنه يقيس أكثر من سمة من السمات، وبالتالي يمكن أن تستخدم الاختبار الواحد للكشف عن أكثر من سمة تبعاً لطريقة معالجة الاستجابات التي نحصل عليها.

وقد كانت دراسة فلانجان لاختبار بيرنرويتز بداية لاستعمال درجات السمات كما تحدها القواعد الإحصائية. لقد اعتقد فلانجان أن السمات لا يمكن أن نطلق عليها أسماء خاصة بها إلا إذا كانت الارتباطات فيما بينها منخفضة. وعلى هذا قام بحساب معاملات الارتباط لدرجات اختبار بيرنرويتز من تطبيقه على 305 مراقب، ووجد أن هذه السمات ليست مستقلة. فقد وجد أن الخطوات لا يختلف إلا قليلاً عن النزعة للصبيّة، وعندما طبق التحليل العاملي وجد عاملين هما: الثقة بالنفس والميل الاجتماعي، يمكن أن يسرّ للبيانات التي نحصل عليها من قوائم السمات الأربعة الأصلية، وأنشأ مفاتيح خاصة بقياس العاملين أو المسمّين اللتين وصل إليهما أي ثقة بالنفس والميل الاجتماعي. وقد ارتبطت درجات هاتين المسمّين ارتباطاً ضعيفاً جداً، وعلى هذا أصبح من الممكن أن نقول أنهما يمثلان جانبين مستقلّين من التقرير الذاتي.

بعد أن حدث تطور آخر له مغزاه، ذلك هو إخضاع نظريات للشخصية تماماً لبحث إحصائي عن أبعاد يمكن أن تلخص الشخصية. لقد أدت الارتباطات الدلالية بين عناصر الاختبارات إلى أن اقترح جيلفورد مثلاً، أن الخطوات يمكن أن نقسمه إلى خطوات اجتماعي (ج)، وخطوات فكري (ف)، ولكتّاب (ك)، والقلب العزلي (ت)، وضبط للنفس (ض). وعلى هذا أنشأ الاختبار الذي لسماء: اختبار العوامل (ج - ف - ك - ت - ض).

ثم أضاف بعد هذا ثمانية جواذب أخرى للشخصية. ولم تكن مقاييس جيلفورد مستقلة، ولهذا أصاد بعض الباحثين تنظيم هذه المقاييس في أنماط تجعلها أكثر كفاءة. من هؤلاء مثلاً ترستون الذي استطاع أن يفسر كثيراً من المعلومات في مقاييس جيلفورد الثلاثة عشرة بسبعة عوامل فقط وأعاد تسميتها بحيث أصبحت هكذا للتأمل، الميل الاجتماعي، الاتزان الانفعالي، الحيوية أو الذكورة، السيطرة، النشاط، الاندفاع.

وهكذا نجد أن واحداً من علماء النفس يصنف عناصر الاختبار التي تقاس السمات بشكل دقيق تفصيلي، بينما يأتي التالي ليضعها في مجموعات صغيرة، ثم يأتي آخر فيعيد تقسيم المجموعات بطريقة جديدة، ويصلي كلاً منها أسماء جديدة للعوامل التي يصل إليها، وبهذا لا تنتهي هذه اللعبة، كما يطلق كروبيك، وفي رأيه أنه ما لم ترتبط قوائم السمات بنظرية محددة وبمحات خارجية، فإن عملية الاختبار لا يمكن إلا أن تكون عملية تفضيلية شخصية. وليس هناك اتفاق في الوقت الحاضر بين أصحاب التطويل للعالمي، بالنسبة إلى عدد العوامل التي تم تحديدها بشكل دقيق، أو لفضل تنظيم لها، أو لنسب الأسماء لها.

لما الاتجاه الأحدث في قياس الشخصية بصفة عامة فهو تحديد المفاهيم والتكوينات النظرية على أساس نظرية للشخصية ثم إعداد العناصر في الاختبارات بصورة تؤدي إلى الحصول على معلومات عن هذه المفاهيم والتكوينات. وقد كان للنظرية يوقع عن الشخصية تأثير في الدراسات الأولى عن الانطواء، ولكن هذه النظرية لم تتعد التراجع بعض عناصر اختبارات وضعت لمجرد المحولة. أما الآن فهناك دراسات واسعة عن اختبار مايرز - برجز

والتي شملت عناصره ومفاتيح تصحيحه صراحة من نظرية بونج. وتوجد مقاييس أخرى تمثل هذا الاتجاه مثلها مقاييس التفضيل الشخصي لإدواردز.

والذي يشق من نظرية موري عن الحاجات الخمس ضرورة، ومن أمثلة هذه الحاجات: الإنجاز (التحصيل)، الخضوع، والنظام، والعرض، والاستقلال الذاتي، السيطرة، التآمل للنفس (والمقصود بهذه الحاجة تحليل دوافع ومشاعر الفرد الذاتية وعند الآخرين)، ويتكون المقياس من 210 زوجاً من العبارات، وعلى المفحوص أن يختار عبارة من كل زوج على أنها أكثر تمثيلاً لخصائص شخصيته، ومن أمثلة هذه الأزواج:

- أ- أحب أن أتحادث مع نفسي مع الآخرين.
- ب- أحب أن أعمل بهدف معين حدثته بنفسي.
- ج- أشعر بالاكتمال عندما أفضّل في أمر ما.
- د- أشعر بالاضطرار عندما أتحادث إلى جماعة.

ويهدف هذا المقياس إلى تقدير القدرة التنموية الخمسة عشر متغيراً المرتبطة بالحاجات.

ويمكن تقويم الخمسة عشر درجة المرتبطة بالحاجات على أساس معيار مئوية ونتيجة لطالب وطالبات الجامعة، وهذه للمعيار تقويم على عدد 749 طالباً، و 760 طالبة موزعين على جميع أنحاء الولايات المتحدة، وقد أضيفت معايير مئوية للكبار ينتمون إلى بيئات حضرية وريفية.

لما عن ذلك للمقياس فقد ذكر دليل الاختبار معاملات ثبات لمفهوم الفرعية الخمس عشرة معتمدة على إعادة الاختبار بعد أسبوع ولحد تتراوح بين

0.74، 0.88، بينما معاملات الثبات المعتمدة على التجزئة للنسبة تتراوح بين 0.60، 0.87.

أما بالنسبة إلى صدق مقياس إدواردز، فعلى الرغم من قلة البيانات الواردة عنها في دليل الاختبار ذاته، فإن لهجاتاً كثيرة مستقلة حول صدقه قد نشرت، إلا أنها أعطت نتائج متعارضة، وتذهب تستأري إلى أن مقياس التفصيل للشخصي له مزايا كثيرة إلا أنه يحتاج إلى كمين:

أولاً، مراجعة للتخلص من بعض جوانب الضعف الفنية فيه، وبصفة خاصة فيما يتصل بشكل العناصر وتفسير الدرجة، وبالتالي، دراسات صدق تجري بدقة مستبعدة من أساليب تطويل أنماط الدرجات مناسبة للدرجات هذا المقياس التي لا تتحدد فيها قوة الحاجة بصورة مطلقة ولكن بالنسبة إلى قوة الحاجات الأخرى عند الفرد، وعلى هذا فإن إطار المرجع في هذا النوع من الدرجات هو الفرد ذاته وليس عينة التقنين. وقد يكون شخصان متفقان في درجاتهم على المقياس إلا أنهما يختلفان اختلافاً جوهرياً في القوة المطلقة لحاجتهما.

وحسب ثبات الاختبار بطريقة التجزئة للنسبة، وبتطبيقه على طلاب الجامعة، وكان يتراوح بين 0.34، 0.77 وقد استعمل هذا المقياس في بعض دراسات المقارنة بين الثقافات وغيرها.

وإذا ارتبط في تصميمه بالبحث في نظرية هل سبيل السلوكية، وسوف نعود إليه بعد دراستنا لاختبار مينسوتا للمتعدد الأوجه الذي تولاه اهتماماً خاصاً.

## الاختبار منسباً للمتعدد الأوجه للشخصية

يعتبر هذا الاختبار من أهم الأدوات التي وضعت لقياس الشخصية عن طريق التقرير الذاتي، لهذا سوف نتناوله بشيء من التفصيل. نجد لهذا الاختبار مكاناً بين اختبارات الشخصية تشبه مكان اختبار سترونج بين اختبارات الميول، ذلك لأنه يشارك اختبار سترونج في الأساس التجريبي الذي حسم وفقاً له، وكذلك يشبه من حيث تعرضه لبحث عتيق وطويل من جانب واضعيه. وقد ظهر هذا الاختبار في وقت كانت الحاجة إليه شديدة، وكان الاعتماد عليه كبيراً أثناء وبعد الحرب العالمية الثالثة التي شهدت توسعاً كبيراً في علم النفس المرضي والعلاج النفسي. أي أن نشأته كانت أصلاً نشأة إكلينيكية وهذا مما ساعد على انتشاره، وكذلك ساعد على انتشاره أيضاً بعض الأمور الفنية المبتكرة فيه ومنذ نشره في عام 1940 ونشر دليله في عام 1943 فإنه استلزم هيئاً من البحوث التي استخدمته.

وكان تصميم الاختبار في الأصل على يد عالم نفسي هو ستارك هاتلوي وطبيب نفسي هو ج. م. ما كيلي وكان الهدف من هذا التصميم هو المساعدة في تشخيص المرضى النفسيين، وتطوّل القسمات التي تميز لشذوذ النفسي الذي يؤدي إلى المعجز.

يتكون الاختبار من 550 عبارة يسأل المفحوص أن يصنفها وفق أحد القسم ثلاثة، أي يجيب عليها إما بنعم أو لا، ولا يستطيع أن أجزم أو لا أدري. وفي الصورة الفردية للاختبار كانت العبارات تطبع على بطاقات منفصلة وعلى المفحوص أن يصنفها في ثلاث مجموعات، وبعد ذلك طبعت العبارات في كتيب وتسجل الاستجابات في ورقة الإجابة. وقد صممت قصورتان للفردية والجماعية

للجبار، أي من 16 سنة وما بعدها، إلا أنهما استخدما بلجاجة مع مراقبين أسخر من هذا التفسير. وبصفة عامة يفضل استخدام الصورة الفردية وخاصة عدد الأشخاص المضطربين أو من ذوي المستوى التعليمي أو الثقافي المنخفض.

وتتنوع عناصر الاختبار للمتعدد الأوجه للشخصية وتتسع من حيث محتواها فتشمل مجالات مثل: الصحة، والأعراض الجسمية، النفسية، والاضطرابات العصبية، والاتجاهات نحو الجنس، والسياسة، والدين، والمساكن، التربوية والمهنية والعائلية، كما تشمل كثيراً من المظاهر والأعراض السلبية والذهالية المعروفة مثل المخاوف المرضية، والاضطهاد، ونزعات تعذيب الذات وتعذيب الآخرين وغيرها.

وعندما نشر الاختبار في صورته الأولى كان يعطي درجات على تسعة (مقاييس إكلينيكية)، ويتكون كل مقياس منها من عناصر ميزت بين جماعة مرضية محدودة وبين جماعة ضابطة عادية، وعلى هذا فإن هذه المقاييس قد أنشئت بطريقة تجريبية، وذلك عن طريق تقسيم العناصر وفق محك معين وهو التشخيص النفسي المعروف ووفق هذه الطريقة أعدت المقاييس التالية:

- 1- نوع المرض.
- 2- الاكتئاب.
- 3- الهوس.
- 4- الاضطراب النفسي.
- 5- الذكورة، الإكوث.
- 6- البارافيا.
- 7- السيكاثينيا.

8- القياس.

9- القياس الخفيف.

وبما يلي تعريف مختصر بما يدل عليه كل من هذه المقاييس:

- 1- قووم المرض: وهو مقياس للاهتمام للرتد بالوظائف الجسمية، والقلق الذي لا يبرر على الصحة.
- 2- الاكتئاب: ويرتبط هذا المقياس بالشعور بالانقباض والتشاوم والإحساس بحجم الأهمية الذاتية.
- 3- الفهم المرضي: ويقاس هذا المقياس درجة تشابه المفهوم بالمرضى الذين تظهر عندهم أعراض تتخذ صورة شكلي عامة أو محددة مثل العمى أو الشلل أو القيء أو الاضطرابات المعوية أو الأعراض العقلية.
- 4- الانحراف السيكوباثي: وهو مقياس لدرجة تشابه المفهوم بجماعة السيكوباثيين الذين تمثل العنصرية الرئيسية عندهم في نقص الاستجابة الانفعالية السيفة وقلة الاهتمام بمشاعر الآخرين، وعدم المبالاة بالمعايير الاجتماعية.
- 5- الذكورة - الأنوثة: وهو مقياس للذكورة والأنوثة في أنماط الاهتمامات، وقد استخرجت عبارات هذا المقياس نتيجة مقارنة استجابات الذكور من ذوي الاهتمامات الذكورية باستجابات الذكور من ذوي الاهتمامات الأنثوية، أما الفروق بين استجابات الذكور والإناث فقد درسها واضعاً الاختبار إلا أن هذه الفروق لعبت دوراً صغيراً نسبياً في هذا المقياس. وتدل الدرجات المرتفعة على هذا المقياس على غلبة أنماط الاهتمامات للخاصة بالجنس الآخر.



6- **البارتوبا:** وهذا مقياس التشكك الزائد، والحساسية المفرطة، وهواجن الاضطهاد.

7- **السيكاثينيا:** وهو مقياس للخاوف المرضية والملوك القهري، الذي قد يكون صريحاً مثل التكرار الزائد لصل البنين، أو ضمناً كما يتمثل في عدم القدرة على التحرر أو الهروب من فكرة متسلطة.

8- **الفسام:** ويكشف هذا المقياس عن التفكير أو السلوك الخلطي للشاذ والذي لا يتفق مع الاتصال بالواقع وفهمه.

9- **الهُوس القهلي:** وهو مقياس للنشاط الزائد في الفكر والعمل، وعدم التركيز الطويل، والانشغال بعدد كبير من الأمور والمشروعات في وقت واحد.

وتعطى بيانات جماعة الثقلين من العائنين درجة معيارية (ت) يمكن تحويل الدرجات الخام إليها بحيث تضبط لنتائج على بطاقة صحيفة نفسية وتتركز الأهمية الأولى على الدرجات التي تزيد عن 70 (والدرجة 50 هي متوسط جماعة المرجع)، وهذه الدرجة الفاصلة، أي 70، هي درجة مفترقة على أساس قطعي احتمالي، أي ليس لها أساس أو تبرير نظري، وعلى أي حال نجد أن مضمري نتائج الاختبار يعولون إلى الاهتمام بكل الارتفاعات، أو القسم، في بطاقة الصحيفة للنفسية سواء كانت هذه الارتفاعات تتخطى خط (70) درجة أو لا تتخطاه.

#### مقاييس الضبط

ومن العلامات المميزة والخاصة بمقياس مينسوتا المتعدد الأوجه استعماله لأربعة مقاييس تسمى مقاييس التصديق، والتصديق هذا معناه الضبط لأنها

في الواقع تمثل ضوابط للإعمال، وسوء الفهم، والتحلول، ولتكاثر النمط الاستجابات، والاتجاهات نمر أخذ الاختبارات، وهذه المقاييس هي:

1- الأول A: وتمثل درجة هذا المقياس مجموع العناصر التي أجاب عنها المفحوص بالألا أستطيع أن أجزم، ومن المرغوب فيه أن تكون هذه الدرجة أقل ما يمكن.

2- الثاني B: ويمثل الكذب، ويعتمد على مجموعة من العناصر التي تجعل المفحوص يبدو في شكل مقبول اجتماعياً، ولكن ليس من المحتمل أن تكون الإجابة عنها صادقة في الاتجاه المرغوب فيه، ومن أمثلة هذا العنصر (لا أقول الصدق دائماً).

وعلى الرغم من أن الإجابة لصحيفة المعتادة تكون نعم، إلا أن الإجابة المقبولة اجتماعياً تكون لا، ومن المفترض أن الشخص الذي يريد أن يظهر نفسه في صورة مقبولة يحصل على درجة مرتفعة في المقياس B عن طريق تحريف استجاباته لمبارات المقياس، وبمثل الدرجات العالية على هذا المقياس على أن الإجابات لا يمكن الثقة بها، ولكنها لا تدل على أن المفحوص يعتمد الكذب.

3- الثالث C: وهو مقياس النمط، ويتكون من العناصر التي ندر أن أجيب من مجموعة للتقليد في نفس الاتجاه الذي أجابه فيها المفحوص، وعلى الرغم من أن هذه العناصر لا يمكن أن تمثل نمطاً معيناً من الشذوذ النفسي فإنها تعبر عن سلوك غير مرغوب فيه، وعلى هذا فإنه من غير المحتمل في الواقع أن يعبر مفحوص واحد عن كل أو معظم هذه الأعراض. وترتفع الدرجة على هذا المقياس عندما لا يستطيع

المفحوس أن يعطي إجابة مميزة لسبب من الأسباب كل لا يستطيع القراءة أو الفهم بدرجة مناسبة، أو الإعمال، سواء عن قصد أم غير قصد، كما قد يرجع ارتفاع الدرجة إلى الخطأ في التصحيح. وبصفة عامة كلما زادت الدرجة على هذا المقياس زاد احتمال أن بعض العوامل تدخلت لتقلل من صدق الصحيفة للنفسية (أو البروبايل).

4- الفاي K: ويمثل مقياس التصحيح، وفي هذا المقياس استعمال لتركيبية أخرى من العناصر المختارة، وتعطى درجة هذا المقياس دلالة على الاتجاه نحو أخذ الاختبار، وهي متصلة بكلا المقياسين السابقين L، F، ولكن يعتقد على أنها أكثر خفاء منهما وتدل للدرجة العالية على هذا المقياس على اتجاه دفاعي أو على محاولة (للمخادعة في الاتجاه الحسن) أو للتعريف المقصود نحو الطرف السوي، كما تدل للدرجة المنخفضة على الصراحة المفرطة، ونقد الذات، أو المحاولة المتعمدة للمخادعة في الاتجاه السيئ أو الميل إلى إظهار النفس بالمظهر غير السوي.

ومن المهم أن نشير إلى أن استعمال مقياس الضبط الأربعة هذه ليس مقلداً تماماً ولكنه يترك جزئياً لحكم مستعمل الاختبار، وكذلك تشير إلى أن هذه المقاييس تخضع باستمرار لإعادة التحديد والمراجعة في ضوء الأبحاث الجديدة.

كانت تلك هي المقاييس الإكلينيكية للتسعة الأساسية ثم المقاييس الأربعة الأخرى الخاصة بضبط الاختبار، إلا أنه منذ نشر الاختبار أنشئ عند كبير جداً من المقاييس الجديدة بواسطة بلشين آخرين لم يشاركوا في وضع المقاييس الأصلية، من هذه المقاييس الجديدة مقياس عن الانطواء الاجتماعي، وقد أضيف

إلى المقاييس التسعة الأصلية وهو مضمن في الاختبار ورفعه (صفر). ويقاس هذا المقياس للزخعة إلى الانزواء والابتعاد عن الاتصال الاجتماعي بالآخرين.

وبهذا فإن الاختبار في صورته الحالية يعطي (14) درجة على الأربعة عشر مقياس، منها التسعة الأصلية، ومقياس الانطواء الاجتماعي، ثم مقياس المنبسط الأربعة.

ومعايير المجموعة الضابطة الأصلية والتي تتكون من حوالي (700) فرداً مذكورة في صورة درجات ثنائية. أو درجات معيارية بمتوسط (50) وانحراف معياري قدره (10). وتستعمل هذه الدرجات المعيارية في رسم الصحيفة النفسية. وأي درجة تعادل (70) أو تزيد عنها، أي تقع فوق المتوسط بدرجتين انحراف معياري، وتُعتبر إليها بصفة عامة على أنها تتميز انحرافاً مرضياً ومما يجب ملاحظته أن المعنى الإكلينيكي لنفس الدرجة قد يختلف من مقياس إلى مقياس آخر. فمثلاً الدرجة (75) على مقياس توهم المرض وعلى مقياس للفصام لا تدل على نفس الدرجة من حدة الانحراف.

أما تفسير الدرجات المنخفضة على الصحيفة النفسية للاختبار ليس على درجة كافية من الوضوح. إذ يعتقد الإكلينيكيون أن وقوع درجات أقل من (50) كثيراً قد يكون له معنى تشخيصي، ولكن لم توضع بعد تفسيرات منتظمة لهذا.

على أي حال، فإن هناك قديراً من نتائج البحوث التي أجريت على هذا الاختبار يوضح أنه بصفة عامة كلما زاد عدد وحجم الدرجات المتعرفة على الاختبار، زاد احتمال أن يكون لشخص مصاباً باضطراب شديد.

ومن الواضح أن الاستعمال الرئيسي للاختبار المتعدد الأوجه هو التشخيص المتميز وعندما يستعمل في هذا المجال لا بد أن يمارس مستعملة منتهى الجذور ولا يلجأ إلى التفسير الحرفي للمقاييس الإكلينيكية فيه، لأن الأمر أكثر تعقيداً من تلك الأسماء المعطاة للمقاييس الأصلية، فليس معنى الدرجة المرتفعة على مقياس القفصام مثلاً أن الشخص مصاب بالقفصام، فقد يحصل شخص حادي على نفس الدرجة، بينما قد يحصل المصابون بالقفصام فعلاً على درجات مرتفعة على مقاييس أخرى. ولكي يستبعد احتمال سوء التفسير للدرجات على المقاييس الواحد فإن أرقاماً من صفر إلى 9 حلت محل أسماء المقاييس في طبعات الأخيرة من الاختبار. ولكي يسهل النظر في نمط الدرجات أو الصحيفة للنسبة وبصحتها وضع نظام ترميز عددي للصحيفة للنسبة. وفي هذا النظام يدل تتابع وترتيب أرقام المقاييس، بنظرة سريعة، على النقط المرتفعة والمنخفضة في الصحيفة للنسبة لشخص ما. فمثلاً للرمز 2-94 يدل على درجة مرتفعة على المقياس رقم (4) (ب د) وعلى المقياس رقم (9) (م 1) بترتيب متناقص، أي أن الدرجة على المقياس الأول أعلى منها على المقياس الثاني، كما يدل هذا للرمز على درجة منخفضة على المقياس رقم (2) (د).

والواقع أن أسماء المقاييس التي استحدثت كانت تعتمد إلى تصنيف إكلينيكي يقوم على أساس نظري مشكوك فيه على الرغم من شيوخه وقت وضع الاختبار، فالنتصيمات المستعملة والتي وضعت لها مقاييس الاختبار كانت محل شك ومراجعة على حلم النفس المرضي، أما في المجال الطبي الواقعي فإن هذه التصنيفات لم تعد مقبولة.

ومن دواعي الحذر في تفسير درجات المقاييس أن بعض هذه المقاييس لا تتمتع بثبات مرتفع، وتعتبر هذه نقطة ضعف في الاختبار. لأن فاعلية أي تحليل للصحيفة النفسية تقل بسبب أخطاء الصدفة في الدرجات التي تعتمد عليها هذا التحليل، فإذا كانت درجات المقاييس المفردة غير ثابتة ومتداخلة أي ترتبط فيما بينها ترابطاً عالياً فلن كثيراً من الاختلافات فيما بين الدرجات ربما تكون ناتجة عن الصدفة.

وتتراوح معاملات ثبات الاختبار للنتيجة عن إعادة الاختبار، بالنسبة إلى الراشدين الأسوياء وغير الأسوياء بين 0.50 و0.9 وكانت الفترة بين الاختبار وإعادة نتراوح بين بضعة أيام إلى عام. إلا أي هناك دراسة أجريت على طلاب الجامعة كانت معاملات الثبات فيها صوماً أقل من السابقة، كما أظهرت الدراسة نفسها تنوعاً كبيراً بين مقاييس الاختبار كانت تتراوح بين 0.05 لمقياس (ب) 1، 81 لمقياس (ب ت).

وثمة عيب آخر في الاختبار يتصل بمدى تمثيل عينة التقيين، فالدرجات المعيارية التي اشتقت منها التعليمات للصحائف النفسية موضوعة في ضوء جماعة ضابطة تتكون من 700 راشد الذين اختيروا في التقيين الأصلي للاختبار. وتبدو هذه العينة قاصرة جداً إذا قورنت بعينات التقيين الأخرى الأكثر شمولاً والتي استعملت، مثلاً، في تقيين اختبارات القدرة وقد أكتت أبحاث كثيرة ضرورة أن تكون العالير متنوعة تنوعاً كبيراً في جاسات أصل مختلفة، ومن هذه الأبحاث ما وضع أن متوسطات درجات ملبة للكلبات كانت فوق (50) بشكل مطرد على بعض المقاييس، كما بينت دراسة أخرى أن 39% من طلبة

الكليات حصلوا على درجات فوق (70) على مقياس أو أكثر من مقاييس الاختبار.

ويتأثر الاختبار المتعدد الأوجه، مثل سائر اختبارات الشخصية، بالاختلافات الثقافية خاصة بين المجتمعات بل وبالمستويات الثقافية الفرعية داخل المجتمع. وقد أظهرت دراسات كثيرة أجريت في بلاد مختلفة ارتفاعاً ملحوظاً في بعض مقاييس الاختبار عندما وضعت للصحائف للنسبة على المعيار الأمريكية الأصلية. وإن تفسير مثل هذه الاختلافات في ثقافة المجتمع أو ثقافته الفرعية يتطلب معرفة بخصائص هذه الثقافات وطروفها، فقد يكون الاختلاف في درجات الاختبار المتعدد الأوجه راجعاً إلى اختلاف الأفراد في تفسير عبارات الاختبار أو تعليمات الاختبار، فمثلاً قد يرجع الارتفاع في بعض المقاييس عند بعض الجماعات إلى اعتقادات متوارثة نحو التواضع وتقليل شأن الذات. وكذلك تتأثر الدرجة بالاختلافات الثقافية بتصل بالملوك الذي يعتبر مقبولاً لاجتماعياً. إلا أن الدرجات المرتفعة على مقاييس هذا الاختبار، بالنسبة إلى بعض المجتمعات والجماعات، قد تدل على وجود مشكلات لفعالية حقيقية يمكن إرجاعها إلى أساليب تنفذة لطفل، أو الصراعات في الأنوار الاجتماعية، أو غير ذلك من أساليب الاضطراب الانفعالي.

#### الختبارات مشتقة من اختبار الشخصية المتعدد الأوجه

إلى جانب حركة للبحوث للضخمة التي حركها الاختبار المتعدد الأوجه فقد كان له في ذاته خصوصية عالية من حيث اشتقاق مقاييس جديدة منه أو تأسيس اختبارات وتطويرها معتمدة عليه.

فن المقاييس الجديدة للكثيرة التي استخرجت منه عن طريق استصل  
مفاتيح تصبح جديدة لنكر ما يلي:

- 1- مقياس التنصب ويهدف إلى قياس العوامل السيكولوجية المرتبطة  
بالتنصب ضد جماعات الأقلية.
- 2- مقياس المكانة الاجتماعية - الاقتصادية وقيس بعض العوامل  
السيكولوجية المرتبطة بالمكانة الاجتماعية - الاقتصادية للشخص.
- 3- مقياس الاعتراف الجنسي.
- 4- مقياس قوة الأداة.

ويحتوي مقياس تابلور للقلق الظاهر الذي سبقت الإشارة إليه، من  
المقاييس التي اعتمدت على اختبار الشخصية المتعدد الأوجه. وقد استصل هذا  
المقياس استعمالاً واسعاً في البحث وفي الممارسة الإكلينيكية. وقد أنشئ أصلاً  
للاستعمال في التجارب لاختبار فروض معينة تتعلق بتأثير الدافع على التعلم،  
وتحقيقاً لهذا اختبرت عبارات معينة من الاختبار المتعدد الأوجه على أنها تدل  
على تعبير ظاهر عن القلق. وكانت نتيجة هذا الاختبار (50) عبارة، زيد عليها  
عبارات إضافية لا تتصل بالقلق ومختارة من مقياس القسبط للاثلاثة (K, F, L)

كذلك نتج عن الجهود التي بذلت لتعديل الاختبار المتعدد الأوجه بحيث  
يستعمل مع تلاميذ المرحلة الثانوية والطلبة الجدد في الكليات أن أنشئ اختبار  
منصوتا للإرشاد النفسي ويتكون هذا الاختبار من (413) عبارة على المنحصر  
أن يجيب عنها بصواب أو خطأ، ومظم هذه العبارات مأخوذ من اختبار  
الشخصية المتعددة الأوجه. ويغطي اختبار منصوتا للإرشاد النفسي درجات في  
سبعة مجالات هي العلاقات الأسرية، والعلاقات الاجتماعية، والثلثات الانفعالي،



والمسايرة، والتكيف مع الواقع، والمزاج، والقدرة. وعلى الرغم من أن مقياس اختبار الإرشاد النفسي تسمى باسم الطرف المقبول، أي الثبات الانفعالي مثلاً بدل من الطرف الآخر وهو الاضطراب الانفعالي، فإنه تتشابه مع مقياس الاختبار المتعدد الأوجه للمسايرة باسمائها المرضية، فمثلاً مقياس (المسايرة) لاختبار الإرشاد النفسية يشبه إلى حد كبير مقياس الانحراف السيكوباتي في الاختبار المتعدد الأوجه، وكذلك مقياس التكيف مع الواقع يشبه كثيراً مقياس القسما ولهذا الاختبار مقياسان للضبط يشبهان مقياس ضبط في الاختبار المتعدد الأوجه وتشير بعض الدراسات إلى أن مقياس هذا الاختبار ذات قدرة إيجابية ولكنها ضعيفة على تمييز التلاميذ الذين يعرف أنهم غير متوافقين أسرياً من غيرهم ممن يتفقون بتوافق جيد.

ومن الاختبارات التي اعتمدت على الاختبار المتعدد الأوجه اختبار كاليفورنيا للنفسى فهو يشق نصف عباراته من الاختبار المتعدد الأوجه ومن أمثلة مقياسه: السيطرة، التزعة الاجتماعية، ضبط النفس، الإنجاز عن طريق المسايرة، والإنجاز عن طريق الاستقلال والمرونة، وتعتبر أمثالي هذا الاختبار من أفضل اختبارات الشخص المستعملة حالياً، وقد خضع لدراسات وتحسينات عديدة أصلته هذه الكلفة. ولهذا الاختبار صورة لتلاميذ القلوي هي اختبار كاليفورنيا للشخصية.

عرضنا بشيء من التفصيل اختبار الشخصية المتعددة الأوجه لأنه (النموذج الأمثل) لاختبارات الشخصية التي تعتمد على التقرير الذاتي المفصوص، إلا أن الاختبارات التي تنتمي إلى هذا النوع كثيرة ومتنوعة إلى

درجة يكاد يستحيل حصرها جميعاً وخاصة في هذا المجال المحدود، لهذا نختار نماذج ممثلة من هذه الاختبارات التي لم تناقش في الصفحات القليلة الماضية مراعين أن يكون لكل منها ميزته التي تجعله جديراً بهذا العرض الموجز.

قائمة مولي للمشكلات وقد وضعها رومن مولي وزملاؤه، ولها عدة صور من مراحل التعليم المتوسطة حتى الجامعة والرشد، ويحدد فيها المفهوم مشكلاته في أحد عشر مجالاً منها الأخلاق والدين، النواحي المالية والمعيشية، والتوافق للعمل المدرسي، والملاذات الاجتماعية. وتشير الدرجة المرتفعة إلى الحاجة إلى الإرشاد النفسي، وتتخذ العناصر التي يحددها المفهوم أساساً للنقطة الفردية والاجتماعية وقد استخدمت هذه القائمة في عدة دراسات عربية.

#### اختبار التوافق لـ بل

منذ صدور هذا الاختبار في عام 1934 وهو يستعمل على نطاق واسع. والصورة الخاصة بالطلاب مصممة للانتقاء السريع لطلاب المدارس الثانوية والكليات بهدف الإرشاد النفسي، وهذه الصورة تعطي درجات توافق في مجالات أربعة هي: الأسرة، الصحة، الناحية الاجتماعية، والناحية العاطفية. والصورة الأخرى خاصة بالكبار وهي لكل استمالة، وقد زيد عليها مجال خامس هو مجال التوافق المهني.

#### مقياس البورت لدراسة السيطرة والخضوع

يعتبر هذا المقياس من أقوى اختبارات الشخصية المبكرة ويوصف بأنه مقياس للسيطرة والخضوع، لأنه يقيس نزعة الشخص إلى أن يموذ ويسيطر على

زمنائه، أو أن يسلط ويسيطر عليه في علاقات المواجهة اليومية. ويبدأ كل عنصر بوصف موجز لموقف من مواقف الحياة اليومية العادية في اجتماع أو في مدرسة أو سيارة عامة، ويذكر أسلوبين أو أربعة لمواجهة الموقف، وعلى المفحوص أن يحدد أي الأساليب يعبر أكثر عن سلوكه. وتختلف الاستجابات في درجة السيطرة والخضوع التي تمثلها وتعطي لها أوزان مناسبة لهذه الدرجة في التصحيح. وبغلاً عن شيوع هذا الاختبار فإنه قد أُر في كثير من الاختبارات التي نُشئت بعده. ومن المهم أن نشير إلى أن السيطرة كانت إحدى السمات التي برزت بوضوح في دراسات التحلي للعالمي.

### البروفيل الشخصي - جوردن

يختبر هذا الاختبار، وقائمة الشخصية للمؤلف نفسه، من اختبارات الشخصية التي تعتمد على الاختبار الإجباري للتغلب على ما يسمى المرغوبة الاجتماعية، أي ميل المفحوص إلى أن يبدو مقبولاً اجتماعياً، وكذلك الإكلال من الغموض وبعض السمات الأخرى التي تصاحب عملية تصحيح اختبارات الشخصية.

ويقس البروفيل أربعة جوانب للشخصية هي: السيطرة، المسئولية، والارتان الانفعالي، والاجتماعية (أو النزعة الاجتماعية) وهذه الجوانب مستقلة نسبياً، والبروفيل مناسب للاستخدام مع طلاب المدارس الثانوية، والجامعة والراشدين.

ويتكون البروفيل من 18 مجموعة من العبارات الوصفية لشمل كل مجموعة على أربع عبارات وتمثل كل عبارة إحدى سمات الشخصية الأربعة،

المشار إليها من قبل. وتتشابه جملتان من الأربع من حيث أن لها قيمة تفضيلية عالية. أما للجملتان الأخريان فتساويان في القيمة التفضيلية المنخفضة. ويطلب من المفحوص أن يضع علامة على جملة من الجمل الأربعة على أساس أنها تشبهه إلى حد بعيد وعلى جملة أخرى على أساس أنها تشبهه إلى أقل درجة وبهذا لا يستطيع المفحوص أن يستجيب استجابة قبول لجميع الجمل كما يحدث في كثير من مقاييس التثريد لذاتي الأخرى أي أن البروهيل يعتبر أقل قابلية لتشويه من جالاب أولئك الذين يحاولون أن يسلوا صورة طيبة عن أنفسهم.



# اختبارات التحصيل



## اختبارات التحصيل

يقصد بالاختبار التحصيلي الأداة التي تستخدم في قياس المعرفة والفهم والمهارة في مادة دراسية أو تدريبية معينة أو مجموعة من المواد.

وهذه الاختبارات هي أكثر الأدوات شيوعاً في التقويم التربوي، وهي في بلادنا هي الوسيلة الوحيدة التي تستخدم في توجيه التلاميذ وانتقائهم، كما أنها في أغلب الأحوال ما يمكن للمعلم أن يصنعه في ممارسته اليومية للتدريس. ولذلك فإنها تحتل منزلة خاصة بين أدوات التقويم النفسي بجعل من المهم إيراد هذا الفصل لها بحيث يمكن تناولها بشيء من التفصيل حتى يمكن للمعلم والمطلعات التربوية عامة، أن تستفيد من المبادئ التي سوف نتناولها هنا في إعداد هذه الاختبارات.

### تحديد الأهداف التعليمية

أشرنا في الفصل الأول من هذا الكتاب إلى العلاقة بين التقويم النفسي والأهداف التربوية. والواقع أن هذه الأهداف تنبثق من عبارات شاملة التي تدل على الأغراض العامة والقيم والمفاهيم التربوية إلى عبارات المحددة التي تبين أساليب أداء نوعية أو خاصة. ويشار إلى النوع الثاني في علم النفس التربوي المعاصر بالأهداف التعليمية، وإذا كانت الأهداف للتربوية يعجزها الوضوح في أغلب الأحوال، فإن الأهداف التعليمية أكثر تحديداً لأنها تتناول النواتج النهائية لعملية التعلم والتطوير في صورة إدارة يمكن ملاحظته، وبالتالي فهي على درجة أكبر من الوضوح.



ويرى كثير من علماء النفس التربويين المعاصرين أنه من التوجب صياغة الأهداف التعليمية في عبارات صريحة تدل على الأداء النهائي لثلاثة أساليب هي:

1- يؤدي تحديد الأهداف التعليمية تحديداً سلوكياً بهذا الشكل إلى توجيه المعلم في تخطيط عملية التعليم فيمكنه أن يحدد الخطوات التي يجب أن يسير فيها التلميذ للوصول إلى الهدف، كما يمكن أن يلهي بأعباء جميع الاستجابات التي يجب أن تصدر عن التلميذ للوصول إلى الاستجابات الهامة النهائية

ومن المؤكد أن أساليب التدريس تعتمد على الأهداف التعليمية والمخبرات السلوكية، ومع ذلك فإن من نتائج تخطيط المعلم هذا وبخذه قد يكون تحديد الأهداف بتغير مقدار الزمن مثلاً أو المستوى المتوقع للإتيان أو الموضوعات التي يجب تناولها في التدريس، وبالطبع لا يمكن لهذه التغيرات أن تحدث ما لم تتم صياغة الأهداف بصورة صريحة. وفي نفس الوقت لا يستطيع المعلم أن يوائم بين أساليبه في التعليم والفروق في المخبرات السلوكية ما لم يحدد بشكل دقيق ما يستطيع التلميذ أدائه قبل التعليم وما يتوقع منه أدائه بعد التدريس، ومضى ذلك فإن أهم أسباب وجوب صياغة الأهداف بشكل صريح ما يهوه ذلك للمعلم من إرشاد في تخطيط أساليبه في التدريس.

2- بغد تحديد الأهداف التعليمية في عبارات صريحة في تقويم الأداء ومن المعروف أن الاهتمام بالصياغة الدقيقة للأهداف التعليمية بدأ في الأصل العلماء المهتمون ببناء الاختبارات وتقويم المناهج المدرسية، فقد اكتشفوا أن استخدام العبارات الغامضة في التعبير عن الأهداف يجعل عملية بناء الاختبار ومفرداته عملية صعبة إن لم تكن مستحيلة.

3- تحديد ترجمة الأهداف التعليمية في صورة أداء نهائي في توجيه اختيار التلميذ وجهوده.

وقد أثبت ميجر ومكان أن التلميذ يمكن أن يحدث حيث لا يفعل المعلم شيئاً أكثر من إعطاء التلميذ قائمة بالأهداف التعليمية الواضحة بقرم للتلميذ بالاهتمام بها في تعلمهم الذاتي.

### وصف للعمل التربوي

يطلب تحديد الأهداف التعليمية وصفاً دقيقاً للعمل التربوي كما يتطلب تحليلاً تفصيلاً لهذا العمل. وطريقتنا وصف للعمل وتحليل العمل نشأنا أول الأمر في ميدان علم النفس الصناعي ثم انتقلنا في السنوات الأخيرة إلى ميدان علم النفس التربوي.

ومن أهم أساليب وصف العمل التربوي طريقة ميجر وطريقة ميلار، وتتطلب طريقة ميجر في وصف العمل، أو بعجالة أخرى تحديد الأهداف التعليمية خطوات ثلاثة هي:

- 1- تحديد الأداء النهائي الذي يسعى للتعليم أو للتدريس إلى إنتاجه.
- 2- وصف الظروف والأحوال الهامة التي تتوقع حدوث السلوك فيها.
- 3- وصف مستوى الجودة الذي يجب أن يصل إليه أداء التلميذ.

فمثلاً قد يكون الهدف التعليمي الذي نسعى إلى تحقيقه هو أن يستطيع التلميذ تحديد العوامل التي أدت إلى حدوث الحرب العالمية الأولى، وهذا هدف مترجم في أداء يمكن ملاحظته ملاحظة صريحة وعلمة في صورة أداء نهائي

يسمى التحليل إلى إنتاجه (الخطوة الأولى)، ثم يمكن أن نسمي للتلميذ قائمة بالعوامل التي تؤدي إلى الأحداث التاريخية للهامة، وتكون هذه القائمة وصفاً لما يسميه ميجر العوامل للهامة التي يحدث في ظلها السلوك النهائي (الخطوة الثانية)، وأخيراً نتحدد المستويات المطلوبة (الخطوة الثالثة). والواقع أن تحديد المستويات هذا ليس من الأمور للقاطعة ويتراوح بين الإتيان الكامل وبين أقل مستويات الإتيان. وتشير طريقة ميللر في وصف العمل إلى وجود ثلاث خطوات أيضاً هي:

- 1- ظروف الاستثارة أو ما يسميه indication.
- 2- للتنشيط أو الاستجابة التي يجب أن تصدر عن التلميذ.
- 3- للتخذية الراجعة أو تحديد مستوى ملائمة الاستجابة.

والواقع أن هذه الخطوات لا تختلف كثيراً عن تلك التي أشار إليها ميجر، فظروف الاستثارة هي ما يشير إليه ميجر بالعوامل أو للظروف أو الأحوال للهامة، والتنشيط عند ميللر يقابل ما يسميه ميجر بالأداء النهائي، ويتشابه مفهوم التخذية الراجعة عند الأول مع مفهوم مستوى الأداء عند الثاني.

### تحليل العمل للتربوي

بعد ما يقوم للمعلم بوصف الأعمال التي يتوقع من تلاميذه أداء ما يكون عليه تحليل هذه الأعمال وتصنيفها. والفرض الرئيسي من تحليل العمل هو مساعدة المعلم على تحديد أنسب ظروف التعلم، والتي تختلف من نوع من أنواع التعلم الأخر. فخطم تكوين المفاهيم أو حل المشكلات أو الابتكار يتطلب ظروفًا تختلف عن اكتساب المهارات الحركية أو حفظ قائمة من الكلمات. وهذا يعني

بالطبع أن مسؤولية المعلم أكبر من مجرد وصف العمل التربوي أو صياغة الأهداف في عبارات سلوكية صريحة، فهي تشمل أيضاً تصنيف هذه الأساليب السلوكية للمقتضية في كل هدف من هذه الأهداف ثم تحديد أنسب شروط التعلم لكل منها.

ويوجد في الوقت الحاضر تصنيفان أساسيان للسلوك لهما أهميتهما التربوية هما تصنيف جاجليه وتصنيف بلوم. ويشمل تصنيف جاجليه ثمانية أنواع من التعلم تتدرج من البسيط إلى المركب هي:

- 1- التعلم الإشاري أو ما يسمى في سيكولوجية التعلم بالاشتراط للكلاميكي ولديه يتم تعلم استجابة شرطية لمثير أو إشارة معينة.
- 2- تعلم المثير، الاستجابة أو ما يسمى بالاشتراط الإجرائي.
- 3- التمسك أو ما يسمى اكتساب المهارات الحركية (غير اللفظية).
- 4- الارتباط اللفظي.
- 5- التمييز المتحد.
- 6- تعلم المفاهيم.
- 7- تعلم المبادئ.
- 8- حل المشكلة أو التفكير.

لما للتصنيف الذي وضعه بلوم ومساعدوه الميدان العقلي المعرفي فيشتمل على فئتين كبيرتين هما المعرفة والقدرات والمهارات العقلية بحيث نحصل في النهاية على ستة فئات للأهداف التربوية هي:

- 1- المعرفة.

- 2- الفهم.
- 3- التطبيق.
- 4- التحليل.
- 5- التركيب.
- 6- التقييم.

### وصف المداخلات السلوكية

تتطلب عملية التعليم وصفاً للمدخلات السلوكية لدى التلاميذ حتى يمكن المعلم أن يخطط لتنفيذ عملية التدريس، بل وحتى يمكن السلطات التربوية أن تخطط للعملية التربوية عامة ويتطلب هذا الوصف تحديداً دقيقاً للوضع الراهن لمعرفة التعزيز ومعلوماته مهاراته وطرق التفكير وقدراته واستعداداته وميوله ودوافعه وخصائص شخصيته، قبل أن نسمى إلى إكسابه للنماذج السلوكية الجديدة من خلال التعلم والتدريس. وهذه البيانات يجب استخدامها لأغراض الوصف وليس لأغراض التفسير، لأن هذه المفاهيم تصف في الواقع المستوى السلوكي للفرد في لحظة معينة، ولا تستطيع الاختبارات والمقاييس التي نستخدمها في قياس هذه الظواهر أن تفسر لنا أداء الفرد فيها، فلا يمكن مثلاً أن نعرف ضعف الأداء في اختبار الذكاء إلى (ضعف الذكاء) لأن ذلك يصبح دافعاً من الحصول والحاصل ولا يزيد فهمنا للملوك الإنساني.

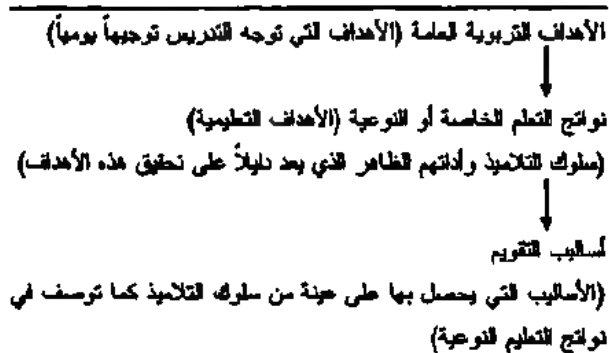
والواقع أن وصف المداخلات السلوكية ضروري لأي برنامج تربوي. فعنصر التعليم يجب أن تبدأ مما لدى التلاميذ ويجب أن نستمر إلى حيث يستطيعون من أداء نهائي تحت شروط معينة والمستوى معين. لما أن تحدد أهدافنا يمكننا أن نقرر ما سوف نعلمه وما نفترض أنه لدى التلاميذ بالفعل.

كن للتفاعل بين الأهداف والمداخلات السلوكية أن يؤدي إلى تغيير الأ<sup>ه</sup>  
 ثبت أنها غير واقعية، كما تتفاعل المداخلات السلوكية مع وسائل التدر  
 إليه، فحين لا تبدأ تدريسنا لتطبيق هدف تربوي معين بمواد تعليمية م  
 أ وإلما ند نستدعي الاستجابات التي سبق تعلمها وخاصة تلك التي نتد  
 موقف للتعلم الجديد.

ثم إن لكل نوع من التعلم مداخلته السلوكية فلتسمية سلوك حل المتد  
 لب معرفة بمداخلات سلوكية تختلف عما يتطلبه لتسمية إحدى المهام  
 يكية.

علاقة بين الأهداف التطويرية وأساليب التقويم التربوي

يلخص جرونلاند العلاقة بين الأهداف التربوية والأهداف المتعل  
 لاليب التقويم التربوي على النحو الآتي:



ويطلب هذا ربط أساليب التقويم مباشرة بنواتج التعلم التي نقومها وبهذا نتيق من أننا نقوم بالتلميذ نحو الأهداف التي أخذناها كإطار لعملية التدريس وعملية ربط أساليب التقويم بنواتج التعلم النوعية هي في أساسها عملية تحليل منطقي وحكم. وهذه العملية يمكن تيسيرها باستخدام خطة ما من خطط التقويم المنظمة، هذه الخطة يجب أن تشمل على الأقل قائمة بنواتج التعلم المرغوب فيها والأساليب التي نستخدم في تقويم التقدم لحواها. وإليك خطة على مبدل المثال في ميدان علم الأحياء للمصف الأول الثانوي:

### الأهداف التعليمية ونواتج التعلم النوعية، أساليب التقويم:

1- يعرف التلميذ المصطلحات الشائعة المستخدمة في علم الأحياء حينما:

- يعرف المصطلحات الشائعة في علم الأحياء.
- يميز بين المصطلحات الشائعة على أساس المعنى.
- يتعرف على معاني المصطلحات الشائعة حين نستخدم في نص معين.

2- يظهر التلميذ القدرة على التفكير الناقد حينما:

- تميز بين الحقائق والآراء.
- يشق استنتاجات صحيحة من المعطيات.
- يتعرف على الافتراضات وراء الاستنتاجات.
- يتعرف على تولحي القصور في المعطيات.

3- يؤدي التلميذ الأعمال الأساسية للتشريح بمهارة حينما:

- يضع العينة في الوضع الصحيح للتشريح.

- يقطع بمهارة دون إفساد العينة التي يدرسها.
- يقسم أجزاء البنية في العينة دون إفسادها.
- ينتهي من التشريح في الوقت المحدد.

4- يظهر التلميذ القدرة على تحديد مواضع المعلومات البيولوجية حينما:

- يستخدم كتلوج المكتبة في تحديد للمراجع.
- يعرف المصادر المشتركة للمعلومات البيولوجية.
- يستخدم جداول المحتوى والفهرس عند البحث عن المعلومات في الكتب
- يحدد مدى ملائمة المعلومات لمشكلة محددة.

5- يكون لدى التلميذ اتجاه على نحو الظواهر البيولوجية حينما:

- يوجّل الحكم حتى تتوافر جميع الحقائق أو مضمونها.
- يبحث عن علاقات السبب والأثر في ظواهر البيولوجية.
- يظهر رغبة في التقامير الجديدة للبيانات البيولوجية.
- يفسر للبيانات البيولوجية متحرراً من التحيز.
- يظهر ثقة في للبيانات البيولوجية التي يحمل عليها بالأساليب العلمية.

من هذا الجدول نجد ما يأتي:

1- نواتج التعلم النوعية كما تصاغ في صورة سلوكه للتلاميذ عديدة ومتنوعة بحيث لا توجد أداة تقويم واحدة تغطي تليلاً مباشراً على تحصيل هذه الأهداف التربوية. ولذلك نحتاج إلى جانب الاختبارات الموضوعية إلى قرائن الملاحظة والتقارير القصصية واختبارات المقال.



2- تظهر أهمية الصياغة الواضحة للأهداف ونواتج التعلم من انتقاء أداء التقييم. وفي الواقع حينما تصاغ نواتج التعلم في ضوء سلوك للتلميذ فإنها لا تدل فقط على ماذا نقوم ولكنها ترحي أيضاً بكيف نقوم.

مثال:

(يحدد المصطلحات الشائعة) يوضح لك أسلوب التقييم الذي يجب أن يستخدمه. فهي تدل على أن التلميذ يجب أن يعطي للتعريف بنفسه وعلى ذلك فإن لاختبار المثال من نوع الإجابة القصيرة حيث يعطى للتلاميذ عدة مصطلحات ويطلب منه تعريفها هو الأسلوب الأكثر ملاءمة في تقييم هذا الهدف. أما الاختبار الموضوعي مثلاً من نوع الاختيار من متعدد حيث يطلب من التلاميذ أن يتعرفوا فقط على التعريف لا تتناسب مع تقييم هذا الناتج التعليمي في سيفته هذه، وإلا فإن السيفته تصبح (يتعرف التلميذ على معنى المصطلحات الشائعة)، وفي هذه الحالة يمكن استخدام الاختبار الموضوعي. ويتوقف هذا على عرضنا من التقييم، فإذا كنا نريد أن نتأكد من أن التلاميذ يعرفون المصطلحات عن طريق تحديدهم وتعريفهم لها فإن الإجراء الوحيد الملائم هو أن يطلب من التلاميذ أن يعرفوا مصطلحاتهم. ولقدرة على التعرف على التعريف الصحيح لا يمكن اعتبارها دليلاً على قدرة التلميذ على إعطاء التعريف الصحيح ما لم يكن هناك دليل من البحث العلمي يؤكد أن نوعي السلوك مترتبان ارتباطاً عالياً. وحيث أنه لا يوجد مثل هذا الدليل بعد فإن الإجراء الأحرط هو اختبار أسلوب التقييم الذي يرتبط مباشرة بناتج التعلم النوعي (الهدف التعليمي).

## تحليل محتوى المادة الدراسية

أشرنا إلى أن الاختبار التحصيلي يرتبط بمادة دراسية معينة ثم ندرسها بالفعل، ومعنى ذلك أنه لا يوجد ما يبرر إعداد اختبارات من هذا القبيل تتناول مواد لم يتم تدريسها أو لم تكون محتوية لمحتوى لعملية التعليم. ولذلك لا بد في إعداد الاختبار التحصيلي من تمثيل كل ما حظي باهتماما أثناء عملية التدريس. ومن الطرق الملائمة في تحديد الأصل الإحصائي المناسب الذي تشتق منه عينة السلوك التي تسمى الاختبار التحصيلي أن تعد قائمة بالموضوعات التي تتضمنها هذه المادة بأكثر قدر من التفصيل وفي أبسط صورة ممكنة. ويمكن للباحث في هذا المسند أن يستفيد بمنهج البحث المعروف باسم (تحليل للمحتوى).

## تحديد الوزن النسبي للأهداف التعليمية ومحتوى المادة الدراسية

بعد تحليل المحتوى يقوم الباحث بتحديد الوزن النسب لكل موضوع (أو وحدة) من موضوعاته، ومن المحركات المستخدمة في تحديد وزن كل موضوع من موضوعات المحتوى هو الزمن المخصص للتدريس كل منها، إلا أنه في بعض الأحيان لا يكون الوقت المستغرق في تدريس موضوع معين دلالة على أهميته، ولذلك فإن الأوزان النسبية للموضوعات لا بد أن تعكس اهتمام المعلم به. وبالمثل لا بد من المعلم أن يحدد الأوزان النسبية للأهداف التعليمية لمادته. فإذا كان يؤكد في تدريسه وصول التلاميذ إلى القسيمات وتطبيقها مثلاً فيجب أن يعطى هذا الهدف وزناً كبير. وبعبارة موجزة لا بد أن تعطي لموضوعات المحتوى والأهداف التعليمية أوزاناً نسبية تعكس الاهتمام الذي تحظى به في عملية التدريس.

ولي تحديد هذه الأوزان قد يعتمد المعلم على خبرته في تدريس المادة وتقبله لمحتواها فيضع قائمة بأهم الموضوعات ثم يصنفها حسب الأهداف التي يرى أن بقيسها الاختبار. إلا أنه في كثير من الأحيان لا يكتفي الباحث بالاعتماد على خبرته الشخصية وحدها ويلجأ إلى بعض الخبراء لاستشارتهم في تحديد هذه الأوزان النسبية.

### إعداد جدول المواصفات

على أساس المطومات التي يتم تصويبها يتم تحديد مواصفات الاختبار والتي تتضمن الموضوعات التي يجب أن يشملها الاختبار ونواتج التعلم التي يجب اختبارها (أي الأهداف التعليمية)، والأهمية النسبية (أو الوزن النسبي) للموضوعات والأهداف. وعلى هذا الأساس يتم تحديد عدد الأسئلة التي يتم إعدادها مرتبطة بكل هدف في كل موضوع والطريقة الأكثر سهولة في تحديد مثل هذه المواصفات هي إعداد جدول ثنائي التصنيف يسمى جدول المواصفات توضع فيه الأهداف التعليمية للتحقق وموضوعات المحتوى رأسياً. والجدول الآتي مثال على ذلك، وملة يوضح عدد المفردات التي يشملها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى في علاقته بكل هدف من الأهداف التعليمية.

المحتوى	معرفة المصطلحات الشائعة	معرفة الحقائق النوعية	لهم المبادئ والتعميمات	تطبيق المبادئ والتعميمات	تفسير الصور والرسوم	لمجموع النسبة %
الطعام	2	6	2	-	-	10
الملابس	2	6	2	-	-	10
النفال	4	2	2	2	5	15
الاتصال	4	2	2	2	5	15
المساكن	-	-	5	5	-	10
حياة المدينة	4	2	6	8	-	20
حياة القرية	4	2	6	8	-	20
المجموع النسبة %	20	20	25	25	10	140

وإذا مثلت خانات هذا الجدول على أساس للوزن النسبي لكل هدف (دورة) للنسب المئوية الموجودة في المظهر الأخير من الجدول، والوزن المعطى لـ موضوع من موضوعات المحتوى (في صورة النسب المئوية الموجودة في مظهر الأخير من الجدول). وكل رقم في كل خانة يدل على النسبة المئوية لأسئلة الاختبار التي تخصص لكل ميدان من ميدان المحتوى وفي كل هدف. أهداف تعليمية. مثلاً بالنسبة لموضوع الطعام تتطرق 2% من أسئلة الاختبار لمعرفة المصطلحات الشائعة فيه، 6% لمعرفة الحقائق النوعية، 2% للمبادئ والتعميمات، وتلك الخانتان للخالتان أمام موضوع الطعام على عـ جرد أسئلة فيه حول تطبيق المبادئ والتعميمات وتفسير الصور والرسوم

والوزن النسبي المعطى لكل ميدان من ميادين المحتوى، ولكل نوع من الأهداف في جدول المواصفات يدل على النسب المئوية الكلية للأسئلة المخصصة لكل منها. فمثلاً في العمود الأخير نجد أن 10% من الأسئلة تتعلق بموضوع الطعام، 10% بموضوع الملابس، 15% بالنقل، وهكذا. وبالمثل فإن المسطر الأخير في الجدول يوضح أن 20% لمعرفة الحقائق النوعية، وهكذا. وبالميلح لا بد من تحديد هذه الأوزان للنسبية قبل إعداد المواصفات كما أشرنا.

### تحديد نوع الأسئلة

أشهر أنواع الأسئلة التي تستخدم في ميادين الاختبارات لتصنيفية أسئلة المقال والأسئلة الموضوعية، وأسئلة حل المشكلات (أو الأسئلة للتصيرة). ونعرض فيما يلي عرضاً موجزاً لأنواع الأسئلة الشائعة في التقويم للتربوي الحديث.

### أسئلة الاختيار من متعدد

وتعتبر أكثر الأسئلة للموضوعية شيوعاً، وتقوم بكفاءة شديدة للنواتج البسيطة للتعلم، بل قد تسمح لقول بعض النواتج للتطبيقات المعقدة. ويتكون السؤال من مشكلة (قد تصاغ في صورة سؤال مباشر أو عبارة ناقصة) تسمى الجذر، وقائمة من الحلول المقترحة (قد تشمل كلمات أو أعداداً أو رموزاً أو عبارات) وتسمى البدائل الاختيارية ويطلب من المفحوص قراءة جذر السؤال ولقائمة البدائل واختقاء البديل الصحيح أو الأفضل. والبديل الصحيح في كل سؤال يسمى ببساطة الإجابة بينما تسمى البدائل الاختيارية المحولات. وهي تسمى

كذلك من واقع وظائفها المقصودة، أي تحويل هؤلاء الذي في شك من الإجابة الصحيحة.

ويستخدم هذا النوع من الأسئلة في قياس كثير من الأهداف التعليمية وخاصة نواتج التعلم في ميدان التفكير واكتساب المعلومات، ومن ذلك معرفة المصطلحات ومعرفة الحقائق للتوصية، ومعرفة المبادئ والقوانين، ومعرفة الطرق والإجراءات، كما تستخدم في قياس الفهم مثل القدرة على تطبيق الحقائق والمبادئ، والقدرة على تفسير علاقات السبب والأثر، والقدرة على تبرير الطرق والإجراءات.

ويتميز هذا النوع بمرورته الشديدة، ومحررة من تقاليد الأنواع الأخرى من الأسئلة، ومن مشكلة أسلوب الاستجابة. بل إن صياغة البدائل الاختيارية قد تؤدي لاستخدامها في أغراض التشخيص. ومع ذلك فإن لأسئلة الاختيار من متعدد مشكلاتها وأهمها أنها تقتصر على المستوى اللفظي للأداء (وهي في هذا تشترك مع غيرها من أسئلة الورقة والقلم)، كما أنها لا تصلح لقياس بعض مهارات حل المشكلة في الرياضيات والعلوم، والقدرة على تنظيم الأفكار (وهي في هذا تشترك مع غيرها من أسئلة التعرف على الإجابة). ثم إن لأسئلة الاختيار من متعدد نقصة خاصة بها لا تشترك فيها مع غيرها، وهي صعوبة إيجاد عدد كاف من المحاولات المجددة وغير الصحيحة في وقت واحد. إلا أن هذه الصعوبة تتغلب مع زيادة الخبرة في بناء مثل هذه الأسئلة.

### أسئلة التفسير

يتطلب هذا النوع من الأسئلة ما هو أكثر من الحفظ واكتساب المعلومات الجزئية، أي أنه يستخدم في قياس نواتج التعلم المرتبطة بالمسلات العقلية العليا كالفهم ومهارات التفكير وفكرات حل المشكلة. ويتكون السؤال

التفسيري من سلسلة من الأسئلة الموضوعية تعتمد على مجموعة مشتركة من البيانات الأولية (المعطيات). وقد تكون هذه المعطيات في صورة مواد مكتوبة أو جدول أو رسوم أو أشكال أو خرائط أو صور. وقد تتخذ الأسئلة المرتبطة صوراً مختلفة ولكنها في الغالب تأخذ صورة الاختيار من متعدد.

### أسئلة المزوجة

تتكون أسئلة المزوجة التقليدية من عشرين متوازيين يحتوي كل منهما على مجموعة من العبارات أو الرموز أو الكلمات. وتسمى المفردات في العمود الذي ننشده له المزوجة بالمقدمات، وتسمى المفردات التي نختار منها الاستجابات response.

وعادة ما تكون أسس مزوجة الاستجابات المقدمات واضحة في ذاتها، وفي بعض الأحيان تتطلب توضيحاً في التعليمات. ويستخدم هذا النوع في قياس الحقائق التي تستند إلى الدعاي البسيط مثل الشخصيات والإنجازات، للتواريخ والأحداث التاريخية، للمصطلحات والتعريف، للقواعد والأمثلة، الرموز والمفاهيم، المؤلفون ومؤلفاتهم، للكلمات الأجنبية ومقابلاتها بالعربية، الآلات واستخداماتها، الثبات وتصنيفاته، الحيوان وتصنيفه، الأشياء وأسمائها، الأعضاء ووظائفها. وتتميز بأنها تقيس بدقة وسرعة مقداراً كبيراً من الحقائق المرتبطة في زمن قصير نسبياً، بالإضافة إلى أنها سهلة الصنع. ومع ذلك فإن لهذا النوع مشاكله وأهمها أنه يقتصر على قياس الحقائق التي تعتمد على التفكير والدعاي البسيط. وكذلك يصعب على الباحث إيجاد مادة متجانسة لها أهميتها في المحتوى الدراسي يمكن أن تصاغ منها قائمة المقدمات.

## مسئلة الإجابة القصيرة (الاستدعاء)

يتطلب هذا النوع من الأسئلة الذي يعرف في بعض المؤلفات باسم مسئلة للتكملة - أن ينتج المفحوص استجابته - وليس مجرد التعرف عليها كما هو الحال في الأنواع السابقة. ويمكن القول أن هذا النوع قد يتطلب إجابة قصيرة إذا عرضت المسئلة في صورة سؤال مباشر، أو تكملة إذا عرضت في صورة عبارة ناقصة. وتوجد أنواع أخرى من هذا النوع منها: مسئلة إعداد القوائم والتي تسمى أحياناً باختبارات المقل ذات التعليمات المحددة و مسئلة التماثل أو التناوب، ومسئلة المشكلات، ومسئلة التعمين (كل ما يتطلب من التلميذ تعيين أجزاء رسم نبات أو حيوان). ويستخدم هذا النوع في قياس الدوافع البسيطة نسبياً للتعلم مثل معرفة المصطلحات، ومعرفة الحقائق النوعية، ومعرفة المبادئ، ومعرفة الطرق والإجراءات، والتفسيرات البسيطة، والقدرة على حل المسائل في الرياضيات والعلوم. وهذا النوع ربما يسبب بساطة نتائج التعلم التي تنمىها (الاستدعاء في أغلب الأحوال)، بالإضافة إلى تحرره من أثر التعمين السائد في اختبارات التعرف بوجه عام. ومع ذلك قد تكون أكبر مشكلاته شيوعاً صعبة تصحيحه.

## مسئلة الصواب والخطأ (أو مسئلة التمييز)

وتتطلب لاختبار إجابة واحدة من إجابتين كالحكم على العبارة بالصواب أو الخطأ، أو الإجابة على السؤال بنعم أو لا، أو الحكم على العبارة بأنها تدل على رأي أو حقيقة، أو تقدير عبارة بالموافقة أو المعارضة. ويستخدم هذا النوع في قياس نتائج التعلم التمييزي البسيط ومع ذلك فهذه أن تكون في قياس الفهم (كالتعرف على علاقات السبب والآخر). وقد تكون أكثر مشكلات هذا النوع من الأسئلة خطورة تأثرها بالبالغ بالتعمين.



## أسئلة الترتيب

وفيها يقوم المفحوص بإعادة ترتيب خطوات أو أحداث أو إجراءات أو تواريخ في تسلسل طبيعي أو منطقي.

## أسئلة المقال

رغم الاستخدام الواسع للأسئلة الموضوعية التي أشرنا إليها لا زالت توجد لتلخيص هامة للتعليم لا يصلح لها إلا سؤال، للمقال، ومنها القدرة على عرض وتنظيم وتكامل الأفكار، والقدرة على التعبير الكتابي، والقدرة على إعطاء التفسيرات والتطبيقات للمعلومات لا مجرد استعراضها أو التعرف عليها.

ومن أهم خصائص أسئلة المقال حرية الاستجابة، وفي هذا جدواها كمنهج للتفكير المنطقي، وفيه أيضاً تكمن صعوبات التصحيح التي تجعل منها أداة لكل كفاءة وكفاءة في قياس الحقائق والمعلومات. ومعنى ذلك أن أسئلة المقال يجب أن تستخدم حيث لا تصلح الاختبارات للموضوعية.

## الاختبارات القصيرة

تستخدم هذه الاختبارات كوسائل موضوعية لتقدير الكفاءة التي يؤدي بها العمل (كالكتابة على الآلة الكاتبة) أو قيادة السيارات. وتنقسم هذه الاختبارات إلى ثلاث أنواع.

1- اختبارات التعرف: وتتطلب من المفحوص التعرف على الخصائص الأساسية للأداء (كأن تعترف مقطوعة موسيقية ويطلب من المفحوص تحديد الأخطاء أو التناقض في الأداء أو للتفسير)، أو تحديد الأجزاء

التي تتألف منها إحدى الآلات: أو اختبار الآلة أو الجهاز المناسب لصل  
محين، أو تحديد العيول أو تصنيف الأخطاء، أو انتقاء عمل أدبي أو فني  
ممتاز.

ب- الاختبارات التي تتضمن مواقف تشبه المواقف الطبيعية: فهي تهدف إلى  
قياس الأنشطة الأساسية في العمل، وتسمى أحياناً اختبارات النماذج  
المصغرة.

ج- اختبار عينة للعمل: وهو عبارة عن محاولة (مضبوطة) أو (متقنة) في  
الظروف الواقعية للعمل، وتنقسم هذه الاختبارات إلى نوعين أساسيين:  
أولهما الاختبارات التي يسهل فيها التمييز بين الصواب والخطأ في  
الآداء وبالتالي يمكن تصحيحها بسهولة مثل التصويب، والأداء العضلي  
في التربية البدنية، والتجميع الميكانيكي، والكتابة على الآلة الكاتبة.  
والنوع الثاني هو الاختبارات التي تعتمد على حكم المراقبين والملاحظين  
لتقديم الأداء وإعطاء درجة أو رتبة، ومن أمثلة ذلك محاولة سلمي  
إبراهيم على (تقنين) اختبارات العزف على البيانو، ويتطلب هذا النوع  
استخدام مقاييس التقدير أو قوائم للملاحظة.

### الاختبارات الشفهية

الاختبار الشفهي هو مزيج من اختبار المقال والاختبار العملي، وله  
فائدته في دراسة السمات العقلية التي يستخدمها المفحوص في الإجابة على  
أسئلة معينة، ولذلك فهو أداة ناعمة في تشخيص الصعوبات. بل إنه في بعض  
الأحوال هو الأسلوب الأحدث في تقديم المفحوص كما هو الحال في قياس صفات  
التلاميذ (في الصفوف الأولى من المرحلة الابتدائية)، وفي قياس بعض نتائج  
التعلم اللغوي (كالقراءة الجهرية).

## تحديد عدد الأسئلة

يتحدد عدد الأسئلة التي يتكون منها الاختبار بمقدار الزمن الميسر له. وكثير من الاختبارات التحصيلية يتحدد زمنه بحوالي (50) دقيقة (زمن الحصة المدرسية العادية) وبعض جداول الامتحانات تحدد فترات أطول تصل إلى ثلاث ساعات. وعلى وجه الصوم كلما طال الزمن وزاد عدد الأسئلة كان الاختبار أكثر ثباتاً. ومن النادر من الوجهة العملية أن نحد اختباراً يستغرق أكثر من ثلاث ساعات.

ومن المفيد كما أشرنا أن نعتبر مجموعة الأسئلة التي يكون منها الاختبار عينة من أصل افتراضي لكل الأسئلة الممكنة التي يمكن أن تستخدم في الاختبار. فمدرس الصف الخامس الابتدائي قد يحصل على (100) كلمة لاختبار المفردات اللغوية وذلك بانتقاء للكلمة الخامسة من قائمة تحوي على (500) كلمة درست أثناء العام الدراسي. وفي هذه الحالة تصبح الـ (500) كلمة هي الأصل الذي نأخذ منه عينة الاختبار، وهو أصل حقيقي ومحدد المعالم. إلا أنه في معظم الاختبارات لا نجد مثل هذا التحديد الحاسم للأصل الإحصائي العام للأسئلة، فمثلاً لا يوجد حد أقصى لعدد المشكلات والمسائل المختلفة التي يمكن إعدادها في اختبارات الجبر.

وبالطبع كلما اتسع مدى الموضوع أو القدرات التي يهدف الاختبار إلى قياسها زاد عدد الأصل الإحصائي. وحجم هذا الأصل يمثل الحد الأقصى لحجم عينة الاختبار. ومعنى ذلك أن أسئلة الاختبار لا يمكن أن تكون أكثر عدداً من الأصل التي تشق منه، إلا أن حجم هذا الأصل لا يصنع حداً أدنى لحجم العينة. فالأصل الذي يتألف من (1000) سؤال يمكن أن تكون العينة المشتقة منه (10) أسئلة أو (50) سؤالاً أو (200) سؤال ولكما زاد حجم الأصل قل تجانسه، أي أصبح يحتوي ميادين مختلفة وشبه مستقلة من المعرفة والقدرة. وللوصول إلى

نتائج دقيقة يجب أن تكون عينة الاختبار المشتقة من أصل غير متجانس كبير نسبياً من تلك التي تشتق من أصل متجانس. ومع ذلك فلا توجد قاعدة علمية تحدد العلاقة من حجم الأصل والعينة.

ويوجد في الوقت الحاضر اتجاه يؤكد استخدام عدد قليل وكاف من الأسئلة بحيث يغطي معظم المفهومين وقتاً وحاولون فيه جميع الأسئلة بمعدلاتهم الخاصة العالية. والموضوع هذا الاتجاه ثلاثة أسباب:

- 1- أن معرفة الاستجابة ليست هدفاً أولاً للتعلم في معظم المقررات الدراسية وبالتالي لا تعد مؤشراً صادقاً للتصصيل لأنه في كثير من ميادين التصصيل لا ترتبط السرعة والنية ارتباطاً عالياً. وبالطبع قد تكون اختبارات السرعة أكثر ملاءمة في بعض الحالات، إلا أن هذه المواقف هي الاستثناء وليس القاعدة.
- 2- أن قلق الاختبار قد يكون على درجة كبيرة من الحدة - حتى في الاختبارات غير المؤنفة - حين يتزايد التأكيد على العمل بسرعة.
- 3- أن الاستخدام للكفاء للاختبار التصصيلي - الذي يصنعه المعلم - يتطلب أن معظم التلاميذ يستجيبون له كله حتى يستخدم نتائجهم في أغراض التمييز.

### تقويم التصصيل في القراءة في المدرسة الابتدائية

قد تكون القدرة على قراءة أهم القدرات التي تؤكدها المدرسة الابتدائية. وفي الصفوف الأولى من المرحلة الابتدائية يحل تعليم القراءة كمهارة أساسية الجزء الأكبر من العمل المدرسي، ومع انتقال التلميذ إلى الصفوف العليا تتفرع دراسته إلى مجالات متنوعة ويقل الوقت المخصص لتعليم القراءة، ومع ذلك فإن لها أهميتها طوال المرحلة الابتدائية. وبعد ما يتقن التلميذ بعض المهارات الأساسية في القراءة، قد تتعمل تحديات طفيفة أهداف تدريس

القراءة بحيث تؤكد للتغيرات الناتجة في المهارات التي نقيسها. ومن ناحية أخرى فإنه لا زال من الأهداف الأساسية للتعليم فهم معاني الكلمات وفهم الجمل والفقرات الطويلة، وهي جميعاً لها أهميتها في جميع للمستويات، ومطابقاً لهذا فإن أي المقترحات تتعلق بتعزيز أهداف القراءة في أحد صفوف المرحلة ينطبق على كل المرحلة. ولذا يلي مثال لأهداف نوعية لتعليم القراءة في الصف الأول الابتدائي:

الأهداف النوعية	الهدف المرحلي (في نصف العام)	فرض الحلم
1- أن يستطيع الطفل أن ينطق بكلمات مثل مدرسة، الحقيقة، ورد، حينما يراها مكتوبة.	أ- أن يستطيع التلميذ قراءة القسم الأول من كتاب القراءة بسرعة وفهم.	أن يستطيع الفرد المتعلم قراءة اللغة القومية قراءة فعالة
2- أن يستطيع الطفل أن يربط كلمات مثل ولد، طيارة، بيت وبين الصور التي تعبر عنها.	ب- أن يستطيع الطفل	
3- أن يستطيع الطفل أن يجيب على أسئلة تتول حول مواد جديدة يقرأها من مختارات من كتاب القراءة.	القراءة الجهرية بسهولة مع بعض التفسير.	
4- أن يستطيع الطفل أن يميز بين الجملة الخبرية والجملة الاستهامية.		

ومن الواضح أنه لو كان الوجب هو تقويم هذه الأهداف النوعية جسيماً فيجب أن تستخدم أساليب مختلفة. فلهذه (1، 4) يمكن تقويمها ببعض صور الأداء الشفوي. بينما الهدفان (2، 3) يمكن تقويمها بهذه الطريقة أو باستخدام بعض المواد المكتوبة. وبسبب الفوائد الكثيرة الخاصة باستخدام الاختبارات الموضوعية في المدرسة الابتدائية فإن على المعلم أن يختار لقياس تحصيل الهدف النوعي رقم (2) طريقة المزاوجة بين صور الأشياء والكلمات التي تدل عليها وفي هذا المستوى يجب أن يعطى لجميع التلاميذ تعليمات شفوية تقرر أنهم يجب عليهم أن يرسموا خطأ يصل بين الكلمة والصورة التي تدل على معنى الكلمة. ومثل هذا السؤال يعطى قياساً مباشر أقرء الطفل على ربط كل كلمة بالصورة الملائمة. والعلاقة البسيطة بين الهدف وهذا الاختبار تدل على فائدة صياغة الأهداف في عبارات تدل على سلوك يمكن ملاحظته.

وبالنسبة لتقويم الهدف (3) فإن المعلم قد يطلب من تلاميذه الانتقال إلى مواد جديدة وقراءتها قراءة صامتة. وبعد هذه القراءة توجه إليهم أسئلة يطلب منهم كتابة إجاباتها عليها وتتعلق هذه الأسئلة بمستوى الفقرة. وهذا الاختبار قد يأخذ صورة أن يطلب من التلاميذ أن يكتبوا جملة في الإجابة على السؤال، وهو صورة مبسطة لاختبار للعقل، إلا أن التلاميذ في هذا المستوى لا تتوقع لهم أن تكون عندهم سهولة كبيرة في الكتابة، ولذلك فإن بعض الاختبارات الموضوعية ينتج قياساً أكثر صدقاً لفهم القراءة عندهم مثل أسئلة الاختبار من متحد مثل:

يوم العيد ذهبنا إلى: الحقل

المدرسة

الحديقة

ومرة أخرى يجب أن تكون تعليمات هذه الأسئلة شفوية وتطلب من التلميذ أن يضع خطأً أو يرسم دائرة حول الإجابة الصحيحة للسؤال.

وحيث أن القراءة هي مهارة لتصال، فإن جميع أهداف للتعليم في هذا الموضوع تهتم أساساً بالفهم، وخاصة فهم معاني الكلمات، فمن المتوقع من التلميذ أن يشرح معنى كلمة، أو يختار مرادفاً لها، أو يربط بينها وبين الصورة أو الشيء الذي تصفه، كما تربط الأهداف الأخرى بينهم الجمل أو الفقرات.

ومع تقدم الطفل في صفوف المرحلة الابتدائية فإن التأكيد على التحصيل في القراءة يتزايد اهتمامه بالملمط مختلفة وجوانب مختلفة من الفهم، فتتطلب اختبارات الفهم من التلميذ أكثر من مجرد استدعاء أجزاء من معلومات الحقائق والتفاصيل.

### تطوير التحصيل في الحساب

تتزايد أهمية للمهارات الحسابية والفهم الحسابي في الحضارة المعاصرة تزايداً مستمراً. ومن المهم للتلاميذ المرحلة الابتدائية أن يتقنوا تعليماً جيداً في المهارات الأساسية في الحساب وفي عناصر التفكير الكمي. وللمساعدة في تحقيق هذا الهدف من المهم للمعلم أن يقوم بتقويم مستمر لتقديم كل تلميذ، وهذا مهم على وجه الخصوص في الحساب حيث كل مهارة لاحقة تعتمد على المهارة السابقة، وحيث يفضل للتلميذ إذا فشل في إلقاء أي من المعلومات الأساسية مثل:

الاهداف النوعية	الهدف المرحلي (في نصف العام)	فرض العام
1- أن يستطيع التلميذ أن يعد ويسجل للسند الصحيح لأشياء أو حروف توضعها للصور في حدود (10) وحدات.	أ- أن يستطيع التلميذ أن يعد حتى الرقم (100)	أن يستطيع التلميذ أن يعد حتى الرقم (100) مشكلات العدد والحساب
2- أن يستطيع الطفل العد إلى الرقم (100) بزيادة واحد.	ب- أن يعرف التلميذ حقائق الجمع الأساسية	
3- أن يستطيع التلميذ الإجابة على مسائل الجمع لأزواج من الأعداد التي تتكون من رقم واحد بحيث يكون مجموعها أقل من (10).	ج- أن يستطيع التلميذ فهم المصطلحات الحسابية الخاصة	
4- أن يستطيع التلميذ شرح معنى مصطلحات مثل جمع وطرح ومجموع وأكثر وأقل.	5- أن يستطيع التلميذ أن يختار اختصاراً صحيحاً الأشياء التي تتحدث بمصطلحات مثل الأول والثاني والثالث... إلخ	

وجميع هذه الأهداف النوعية الخمسة يمكن تقويمها إما شعبياً أو تحريراً، فالهدف الأول يمكن قياسه باستخدام أسئلة بسيطة يعطى فيها للطفل أعداداً مختلفة من الأشياء عليه أن يعدها.



تعليمات (شفوية): اكتب على الخط الموضوع أسهل كل صورة العدد الذي يدل على عدد الحروف في كل منها:

د	د	ص	ص	ح	ح	5	5	5
د	ص	ص	ص	ح	ح	5	5	
	ص	ص		ح	ح	5	5	5

ويمكن تقويم الهدف التالي بسؤال التلميذ أن يعد جهرياً أو أن يكتب الأعداد من 1 إلى 100.

والهدف الثالث يمكن قياسه بكفاءة أكثر بمجرد سؤال التلميذ أن يكتب إجابته على عدد من المسائل التي تتضمن الجمع البسيط:

3 7 5 1 6

4+ 2+ 4+ 8+ 3+

والهدف الرابع يمكن تقويم بسؤال التلميذ مباشرة أن يشرح هذه المصطلحات شفوهاً أو تصريحاً. ويوجد تعديل طفيف لهذا الهدف بحيث يتطلب من التلميذ أن يتعرف على صور أشياء أو عمليات يمكن أن نوصف لفظياً كأن يعطى له رسم لعدد من الكرات متألوة الحجم ويطلب منه وضع خط تحت الكرة الأصغر.

والهدف الخامس يمكن تقويمه باستخدام أسئلة مماثلة، كأن يعطى له رسم لعدد من الكرات المتساوية ويطلب منه وضع خط تحت الكرة الرابعة مثلاً.

إن لنواع الأسئلة التي ذكرناها حتى الآن تتلخّص أساساً بقهاس معلومات للتلميذ وتذكره لحقائق خاصة أو مصطلحات نوعية، إلا أن المعلم قد يهتم بأنماط أخرى من القدرات وخاصة تلك التي ذكرناها في تصنيف الأهداف التربوية (العلوم) تحت عنوان مثل الفهم والتطبيق.

مثال لترجمة الأعداد والمفاهيم إحداهما إلى الآخر.

1- ما الذي يعطي نفس الإجابة التي يعطيها  $4 \times 5$  ؟

أ-  $4 + 4 + 4 + 4 + 4$

ب-  $5 + 5 + 5$

ت-  $4 + 5$

ث-  $4 \div 5$

2-  $\frac{3}{4}$  تساوي

أ-  $\frac{6}{8}$       ب-  $\frac{5}{8}$       ج-  $\frac{4}{8}$       د-  $\frac{3}{8}$

ومن أهم الأهداف التعليمية التي نحتاج إلى تحقيقها في اختبارات الحساب هي القدرة على تطبيق معرفة المبادئ والأساليب أو الإجراءات (وتقاس باختبارات المسائل الحسابية من مختلف الأنواع).

## تطوير التحصيل في العلوم

مع زيادة التأكيد على العلم في منهج المدرسة الابتدائية ككل، نجد من الطبيعي أن تسمى المدرسة الابتدائية إلى زيادة مدى ما يتضمنه برنامجها من العلوم. والأهداف التي تسمى تحقيقها يجب أن تكون أكثر من مجرد معرفة حقائق معينة أو مصطلحات أو ح فظ مبادئ وقوانين. وإنما يجب أن تتضمن قدرات مثل القدرة على إجراء تجارب بسيطة، واشتقاق المبادئ والتميمات من البيانات أو إعطاء أمثلة للمبادئ المختلفة وتطبيق مثل هذه المبادئ في شرح النتائج أو التنبؤ بها كما يجب أن تتعلق الأهداف باتجاهات وميول معينة.

وحيث أن أغلب معلمي المدرسة الابتدائية يعوزهم الإعداد للمسلم في العلوم فإنهم يحتاجون إلى المساعدة في تحديد المناط للعلم التي تنتج عن مقرر العلوم تحديداً واضحاً وفي الحصول على أمثلة خاصة من القدرات المختلفة ومن الأساليب المفيدة في هذا الصدد إعداد ولقاء واختبارات جيدة وأمثلة جيدة في الاختبارات، فكثير من القدرات يمكن التقليل عليها بطريقة جيدة عن طريق تطبيقها في مواقف تد أمثلة اختبار أمثلة لها أو مواقف تتحول إلى أمثلة اختبار. ومعرفة الطريقة التي يمكن بها تقوم هذه القدرات يساعد في توضيح المعنى الحقيقي للقدرة عند العلم وتشجعه على تدريسها.

وفيما يلي مثال:

الأهداف النوعية	الهدف المرحلي	فرض العلم
1- أن يستطيع التلميذ أن يذكر أسباب ضرورة المحافظة على مصادرنا من الماء والأرض والنبات والحيوان.	أ- أن يفهم التلميذ أهمية المصادر الطبيعية مثل الماء والأرض والنبات والحيوان.	أولاً: على المواطن المتعلم أن يقر مصادر الأمة.
2- أن يستطيع التلميذ تحديد الطرق المختلفة للمحافظة على المصادر الطبيعية.		
1- أن يستطيع التلميذ أن يتعرف على إسهامات علماء مثل باستير و لنيسون وسوك و ماركولي.	أ- أن يشعر التلميذ ببعض الإسهامات الهامة التي يقدمها العلم في مجال الرفاهية العامة.	ثانياً: يقىس المواطن المتعلم التقدم العلمي بإسهامه في الرفاهية العامة.
1- أن يستطيع التلميذ أن يحل مشكلات العلم بالوصول إلى النتائج من الأدلة والقواعد.	أ- أن يفهم التلميذ الطريقة العلمية في حل المشكلات.	ثالثاً: تتوفر لدى المواطن المتعلم وسائل دفاع ضد الادعاء
2- عندما يطلب من التلميذ أن يتخذ قرارات في مواقف فرضية في (مجالات العلوم وغيرها) فإنه يصل إلى تلك التي تدعمها الشواهد والأدلة.	ب- أن يصل التلميذ إلى قرارات على أساس الأدلة والشواهد سواء في مسائل العلم أو المسائل غير المتعلقة بالعلوم	

إن الهدفين النوعيين المشتقين من الغرض العام الأول يهتمان بمعرفة المبادئ والقطر، ويمكن تقويم الهدف الأول بمجرد سؤال التلميذ أن يذكر قائمة إما شفوياً أو كتابية، بعدد معين من الأسباب التي تتعلق بالمحافظة على كل مصدر من المصادر الطبيعية المذكورة.

ويمكن للهدف الثاني أن يصاغ على نحو يشبه الهدف الأول، وبالتالي يمكن تقويمه بأن نطلب من التلميذ أن يذكر قائمة شفوية ويمكن أيضاً تقويم هذا الهدف باستخدام الاختبارات الموضوعية.

مثال: حتى يستطيع الفلاح أن يحصل من أرضه على محصول وغير عليه أن يزرع في حقله:

- أ- نفس المحصول عاماً بعد عام.
- ب- محصولاً واحداً ثلاث سنوات متوالية ثم محصولاً آخر السنوات الثلاثة التالية.
- ج- محصولاً واحداً لخمس سنوات ثم يليه محصول آخر للسنوات الخمس التالية.

ويمكن صياغة الهدف رقم (1) بحيث يسمح للتلاميذ بانتقاء الأسباب الهامة للمحافظة على المصادر الطبيعية، وعلى هذا النحو يمكن تقويمه باستخدام لحظة الاختبار من متحد. أما الهدف النوعي المعتمد على الغرض العام الثاني

ويمكن اختبارها باستخدام أسئلة بسيطة مثل أسئلة المزاجية بين أنواع الإسهامات التي قام بها مختلف للطعام.

والأهداف المشتقة من الغرض العام لتثالث أوسع في توفيرها بوعاً ما. ويمكن أن نقيّد في ذلك الاختبارات الضمنية واختبارات المقابل، حيث يعرض على التلاميذ مواقف مشكلة معينة ثم يطلب منه إعطاء الاستنتاجات والحلول المناسبة. كما يمكن توفير هذه الأهداف باستخدام أسئلة الاختبار من متعدد من النوع التفسيري كما هو مبين في المثال الآتي:

مثال: قام أربعة تلاميذ بالاشتراك في تجربة لاستنبات الفول كما يلي:

- قام المحترّ بمسح بعض حبوب الفول بقطعة من القطن ثم تركها جافة.
- قامت مها بمسح بعض حبوب الفول بقطعة من القطن ثم تركتها مبللة.
- قام خالد بوضع بعض حبوب الفول في طين جاف وتركها جافة.
- قام أحمد بوضع بعض حبوب الفول في طين مبلل وتركه مبللاً.

ومن نتائج هذه التجربة حاول التلاميذ للفصل الحكم على أي النتائج

الأكثية حول الصحيح:

- أ- تثبت حبوب الفول إذا تركت فوق قطعة من القطن الجاف.
- ب- تثبت الحبوب إذا تركت فوق قطعة من القطن المبلل.
- ج- تثبت الحبوب إذا تركت في الطين الجاف.
- د- تثبت الحبوب إذا تركت في الطين المبلل.

د- تحتاج الحبوب إلى أن تظل مبللة حتى تكبت.

و- تحتاج الحبوب إلى أن تظل جافة حتى تكبت.

ز- تحتاج الحبوب إلى التربة والماء حتى تكبت.

## المراجع

- 1- التقويم النفسى، د. فولاد أبو حطب، د. سيد أحمد عثمان، مكتبة الأنجلو المصرية - القاهرة.
- 2- الاختبارات الاسقاطية، سيد محمد غنيم، هدى عبد الحصيد، النهضة العربية.
- 3- التشخيص النفسى، ج1، سيد محمد غنيم، هدى عبد الحصيد، النهضة العربية.
- 4- التشخيص النفسى، ج2، سيد محمد غنيم، هدى عبد الحصيد، النهضة العربية.









# التقويم النفسي

مجمع البحوث الإسلامية  
مكتبة دار البحوث

0518323



دار البحوث الإسلامية  
مكتبة دار البحوث  
10000 شارع 10000  
10000 شارع 10000



10000 شارع 10000  
10000 شارع 10000



دار البحوث الإسلامية  
مكتبة دار البحوث  
10000 شارع 10000  
10000 شارع 10000